DOCUMENT RESUME

ED 065 545 TM 001 659

TITLE	Educational Change: Implications for Measurement. Proceedings of the 1971 Invitational Conference on Testing Problems.
INSTITUTION	Educational Testing Service, Princeton, N.J.
PUB DATE	72
NOTE	157p.
AVAILABLE FROM	Educational Testing Service, Princeton, N.J. ($2.00)
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	*Conferences; Early Childhood Education; Economic Factors; *Educational Change; *Educational Opportunities; Measurement Techniques; Open Education; *Problem Solving; Research Methodology; School Role; Social Development; Social Values; Technical Education; *Testing; Vocational Education

ABSTRACT

Papers of the 32nd Educational Testing Service Invitational Conference on Testing Problems are presented. The conference was divided into three sessions. Session one examines the current educational context and includes the following papers: "Educational Opportunity: The Context and the Reality," "The System-Maintenance Function of Social Development in the Schools," and "Economics and the Values of Society." The second session deals with the measurement context. Papers include: "Increase in Educational Opportunity Through Measurement," "Today's Testing: What Do We Know How to Do?," and "Research Methodology for Educational Change." In the final session, the focus is on three ways of broadening educational opportunity and their implication for testing. Papers include: "Exploring Selective Problems and Approaches to Early Childhood Education," "Broadening Education Through Vocational and Technical Training," and "Enriching Education Through Schools Without Walls." (Author/CK)

ED 065545

EDUCATIONAL CHANGE: IMPLICATIONS FOR MEASUREMENT

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY

ETS

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

TM 001 359

PROCEEDINGS OF THE
1971 INVITATIONAL CONFERENCE
ON TESTING PROBLEMS

ETS
EDUCATIONAL TESTING SERVICE

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Educational Change:
Implications for Measurement

EDUCATIONAL CHANGE: IMPLICATIONS FOR MEASUREMENT

PROCEEDINGS OF THE
1971 INVITATIONAL CONFERENCE
ON TESTING PROBLEMS

EDUCATIONAL TESTING SERVICE
PRINCETON, NEW JERSEY
BERKELEY, CALIFORNIA
EVANSTON, ILLINOIS

The thirty-second Invitational Conference on Testing Problems, sponsored by Educational Testing Service, was held at the New York Hilton, New York City, on October 30, 1971.

Chairman: William E. Coffman,
University of Iowa

PHOTO CREDITS: page xii, courtesy of Stanford University.
All other photographs by Hugh Rogers.

Copyright © 1972 by Educational Testing Service. All rights reserved.
Library of Congress Catalog Number: 47-11220
Printed in the United States of America

ETS Board of Trustees
1971-72

John H. Fischer, *Chairman*
President
Teachers College, Columbia University

George F. Baughman
President
The College Foundation

John Brademas
Member, United States House of Representatives
Third District of Indiana

John T. Caldwell
Chancellor
North Carolina State University

Arland F. Christ-Janer
President
College Entrance Examination Board

Charles C. Cole, Jr.
President
Wilson College

John B. Davis, Jr.
Superintendent of Schools
Minneapolis, Minnesota

Hugh M. Gloster
President
Morehouse College

Roger W. Heyns
President
American Council on Education

Wayne H. Holtzman
President
Hogg Foundation for Mental Health and Hogg Professor of Psychology and Education
The University of Texas

Malcolm C. Moos
President
University of Minnesota

John A. Perkins
Vice President, Business and Finance
University of California

Paul C. Reinert, S.J.
President
St. Louis University

William W. Turnbull
President
Educational Testing Service

Albert N. Whiting
President
North Carolina Central University

Charles E. Young
Chancellor
University of California at Los Angeles

ETS Officers

William W. Turnbull, President
Robert J. Solomon, Executive Vice President
David J. Brodsky, Vice President and Treasurer
Samuel B. Gould, Vice President
John S. Helmick, Vice President
Richard S. Levine, Vice President
Winton H. Manning, Vice President
Samuel J. Messick, Vice President
Charles E. Scholl, Vice President
E. Belvin Williams, Vice President
Bernard L. Tchorni, Assistant Vice President
Catherine G. Sharp, Secretary
Russell W. Martin, Jr., Assistant Treasurer
Herman F. Smith, Assistant Treasurer

ETS President William W. Turnbull opens the thirty-second Invitational Conference on Testing Problems.

Foreword

Every year we gather for the Invitational Conference on Testing Problems, I am struck by two things. One is certain similarities: we see many of the same welcome faces from other years and a tackling of persisting and familiar problems. There are also the differences, and at times they are quite striking.

This year seems to me an important case in point. I sense that this conference may mark the beginning of what might be termed a new synthesis in testing.

In recent years we have often heard the concerns of our critics and the critics of American education. Much of this criticism has been legitimate, and good for us to have to meet. Now we are at the point where we must go beyond criticism, not merely with piecemeal answers—not merely with the launching of an excellent new program here or there—but by pulling together some larger consolidation of the answers.

Over the past decade, we have gone from great aspirations in education into a time of disillusion with aspiration in both education and evaluation. The Invitational Conferences on Testing Problems have reflected this experience. But they have also reflected a steady movement not simply toward the mastery of new measures, but toward a new attitude concerning measurement—primarily, a conviction that we must not only bring education and measurement more in tune with each other, but that we must insist that both become much more attuned to the real needs of our time.

And so one may hope that we may now be on the verge of a consolidation of these gains of the past decade—a consolidation of the results of a pervasive attitudinal change—that will bring about a new synthesis in testing.

From knowing William Coffman over many years, I can think of no better person to have helped us gather to move in this direction. And the papers of this conference do much to propel us, if not to the millennium, at least toward something much better than we have known before.

William W. Turnbull
President

Conference Chairman William E. Coffman (right) talks with Harold Gulliksen, ETS.

Preface

The theme of the 1971 Invitational Conference on Testing Problems, *Educational Change: Implications for Measurement*, is one that might have been applied to most of the conferences in recent years. In 1970, speakers looked at the development of large-scale information systems in education and asked how they might be expected to affect the measurement field. In 1969, the spotlight was on the concept of accountability and on the need for a clearer conceptualization of what it is the changing educational system is trying to produce and therefore what should be the focus of the testing effort. In 1968, the papers focused on the challenge to the evaluation specialist posed by the need to monitor large-scale innovative programs supported by the federal government, particularly those aimed at reducing the gap between the educational opportunities for the vast majority of the population and those for disadvantaged minorities. Educational change is all around us, and the impact can be seen in the programs of the Invitational Conference on Testing Problems.

In choosing the theme for the 1971 Conference, however, there was no intention simply to provide variations on themes developed in previous conferences. Rather, this seemed an appropriate time to try to gain some perspective. Critics, both within and outside the measurement field, have been so vocal that there is danger we may conclude that we've been doing nothing right. For example, critics say the schools are rigid and that standardized tests contribute to rigidity. They say that educational standards are low and that norm-referenced tests provide no reference points for establishing standards. They say that multiple-choice tests reinforce the values of convergent rather than of divergent thinking and reduce the probability that creative thinking will be stimulated by the schools. They point to evidence of cultural bias in tests and argue that our evaluation methods set up road blocks to advancement for minority children.

These are serious charges. If true, they require the dedicated attention of professional educators and evaluators. In order to gain perspective, we selected speakers for the 1971 Conference who could take a broad view of what has been going on in education and measurement, of what is now going on, and of what the future is likely to hold.

The first session examines the current educational context. Dean

Goodlad provides a perspective on the changing concept of educational opportunity and concludes that perhaps tests ought to be used to select the less able for schooling and to keep out the more able, who have less need for schooling. Dr. Hess shows how a systematic analysis of the content of textbooks can raise questions about the way the society is portrayed to the student. Dr. Wallace, drawing on her background in economics, throws additional light on the concept of "human capital" introduced to participants in the 1968 Conference by J. Alan Thomas. Dr. Wilson points up the flaw in our historical view of equal educational opportunity but ends on a note of hope with a review of the decision of the California Supreme Court in the case of John Serrano Jr. vs. Ivy Baker Priest. It seems clear that the promise of educational opportunity for all is not yet realized and that it will require the efforts of more than the measurement community alone if it is ever to be realized.

The second session deals with the measurement context. Dr. DuBois provides the historical perspective by asking how the early leaders in measurement viewed their work and concludes that there were as many and varied viewpoints then as now, but that central to all was a desire to improve education through measurement. Dr. Hieronymus examines the current scene and concludes that we can measure whatever we decide is worth measuring, that the problems are philosophical and moral, not technical. Dr. Messick shows how the system concept leads to more complex models for evaluation than are commonly found in today's educational research. Dean Merwin brings the three into a single perspective that can guide evaluation efforts in the '70s.

The luncheon speaker, Dr. Gould, provides a sturdy bridge between the morning sessions devoted to context and the afternoon session devoted to specific innovations. He documents the need for innovation and then outlines tendencies in the culture that threaten it, and he wisely cautions that what is needed is a blending of the good in the old with the new that is good.

In the final session the focus is on three ways of broadening educational opportunity and their implication for testing. Dr. Taylor argues that kindergarten is too late and that the educational frontier is in the period from conception to age three. Dr. Hoyt points to the potential frustration of an educational program that has at center a value system that most students cannot accept as relevant. He points to the need for an educational program with career development at its center and for a wide variety of new measurements. Dr. Watson describes new educational programs that challenge the very foundation of today's norms and present the measurement specialist with challenges sufficient to occupy a lifetime. Dean Allen is impatient with both educators and measurement specialists, and calls for imaginative new educational programs and responsible evaluation of them based on the best evidence available. It is clear that there is plenty to occupy the next generation of measurement people.

The opportunity to preside over such an array of speakers dealing with such a group of challenging topics has been a joy such as must be experienced by one who has climbed a high mountain and looked downward and outward. Without the broad resources of ETS to call upon during the planning stages and the ever-present assistance of Anna Dragositz to keep things moving, I doubt that I could have made the climb.

William E. Coffman
CHAIRMAN

Lee Joseph Cronbach

EDUCATIONAL TESTING SERVICE

Measurement Award

1971

⁂

The ETS Award for Distinguished Service to Measurement was established in 1970, to be presented annually to an individual whose work and career have had a major impact on developments in educational and psychological measurement. The 1971 Award was presented at the Conference by ETS President William W. Turnbull to Professor Lee Joseph Cronbach with the following citation:

> With the daring of the innovator and the rigor of the critic, Lee J. Cronbach, throughout his distinguished career, has constantly probed and ultimately refashioned many of the fundamental principles of educational and psychological measurement.
>
> His systematic investigations of test reliability led him to insights which liberalized and extended that concept into a theory of the generalizability of behavioral measurements. His contributions to the foundations of construct validity helped demonstrate the relationship of test validation to hypothesis testing, thereby placing the validation process in the mainstream of psychological science.
>
> Over the years, Dr. Cronbach's critical analyses have clarified numerous issues—from the statistical treatment of Rorschach scores to the separation of creativity and intelligence, from the analysis of interpersonal perception to the measurement of change. His research has stimulated new areas of activity and controversy, such as the role of response sets in testing and of aptitude-treatment interactions in learning. His innovations in methodology have rechanneled whole fields, as in his applications of decision theory to problems of selection and classification, and his development of strategies for educational evaluation. He has contributed not only to the science of education but also to the development of its scientists through two widely used textbooks.
>
> For his many contributions to the theory, research, and methodology of educational and psychological measurement, Educational Testing Service is pleased to present its 1971 Award for Distinguished Service to Measurement to Lee Joseph Cronbach.

Contents

vii Foreword by William W. Turnbull

ix Preface by William E. Coffman

xiii Presentation of 1971 Measurement Award to Lee J. Cronbach

Session I | The Educational Setting: Historical, Social, and Economic

3 Educational Opportunity: The Context and the Reality, John I. Goodlad, University of California and Institute for Development of Educational Activities, Inc.

15 The System-Maintenance Function of Social Development in the Schools, Robert D. Hess, Stanford University

27 Economics and the Values of Society, Phyllis A. Wallace, Metropolitan Applied Research Center, Inc.

37 Discussion: O. Meredith Wilson, Center for Advanced Study in the Behavioral Sciences

Session II | The Measurement Context: Past, Present, and Future

47 Increase in Educational Opportunity Through Measurement, Philip H. DuBois, Washington University

57 Today's Testing: What Do We Know How to Do? A. N. Hieronymus, University of Iowa

69 Research Methodology for Educational Change, Samuel Messick, Educational Testing Service

83 Discussion: Jack C. Merwin, University of Minnesota

Luncheon Address

87 The Birth of a Non-Traditional Tradition, Samuel B. Gould, Chairman, Commission on Non-Traditional Study

Session III | New Directions in Education: The Challenge to Measurement

101 Exploring Selective Problems and Approaches to Early Childhood Education, Jerome Taylor, University of Pittsburgh

113 Broadening Education Through Vocational and Technical Training, Kenneth B. Hoyt, University of Maryland

125 Enriching Education Through Schools Without Walls, Goodwin Watson, Union for Experimenting Colleges and Universities

133 Discussion: Dwight W. Allen, University of Massachusetts

Session I

The Educational Setting: Historical, Social, and Economic

John I. Goodlad examines the reality of educational opportunity.

Robert D. Hess describes a research project on social studies textbooks.

Educational Opportunity: The Context and the Reality*

John I. Goodlad
Center for Advanced Study in the Behavioral Sciences

Introduction

In this country, the term "educational opportunity" is almost inextricably tied up with "*equal* educational opportunity." This paper discusses both. Likewise, it deals with both education and schooling, two concepts frequently and erroneously used synonymously. Implicitly, the paper recognizes that education is the broader concept of which schooling is a part.

This is not an historical paper in any structural, and especially chronological, sense. Part I deals with the changing concept of educational opportunity. It reveals the difficulty of using any standard yardstick in appraising the status of educational opportunity at one time as compared with another. Part I deals more with the context than the conduct of schooling, whereas Part II focuses on the latter.

Educational opportunity: a shifting concept

"Educational opportunity" is a relative concept. It takes on different meanings with the passage of time, with a shift in place, and when viewed from a socioeconomic, legal, or psychological perspective. Similarly, the concept of "equal educational opportunity" has changed over the years and will change in the future.

---

*The author gratefully acknowledges the bibliographical assistance of Lillian K. Spitzer, Institute for Development of Educational Activities, Inc.

Educational Opportunity: Context and Reality

The meaning of educational opportunity derives in large measure from the *functions* perceived for education in the society. Where upward mobility is virtually closed by a rigid class structure, those in the upper, controlling classes view schools as serving to maintain this structure. Educational opportunity is access to that part of the educational structure designed as appropriate for one's station in life. The system usually consists of tracks or tiers; class status predetermines the appropriate track.

Where class lines are nonexistent or loosely differentiated, the function of education frequently is perceived to be that of providing upward mobility. The function of the schools is to provide at least the minimum core of knowledge and skills requisite to entering and moving up in the socioeconomic system. The schools seek to provide, therefore, a common educational experience for all. The criteria for judging educational opportunity are quantity and availability: how much constitutes a minimum (or, later, adequate) core and how easy is it to gain access to the system?

This last query provides a breeding ground for questions about *equal* educational opportunity; for example, to what extent and on what bases is access difficult for some individuals and groups? Likewise, it is an easy step from preoccupation with quantity to considerations of the quality of educational opportunity. And from input factors such as dollars, teachers, and programs, one moves to output or effects in judging quality. Here, too, the aims or functions to be served by the educational system provide the backdrop for the rhetoric defining the concept. Educational opportunity can be evaluated against the socioeconomic background of access to and retention in the labor market, or against that chimerical screen of individual development.

All of these functions and accompanying perceptions of educational opportunity have characterized the scene in this country. The treatment of Indians and Negroes provides a parallel with the closed system of other places. The expansion of elementary and secondary education has been a response primarily to socioeconomic considerations. But contributing to the G.N.P. is now widely challenged as a criterion for judging educational opportunity.

Much of what follows in Part I derives its structure from Coleman's (1968) excellent analysis of the concept of equal educational opportunity. He points out that, almost from the beginning of the rapid expansion of public education in the nineteenth century, the concept consisted of the following elements:

1. Providing a *free* education up to a given level which constituted the principal entry point to the labor force.
2. Providing a *common curriculum* for all children, regardless of background.
3. Partly by design and partly because of low population density, providing that children from diverse backgrounds attend the *same school.*
4. Providing equality within a given *locality*, since local taxes provided the source of support for schools [p. 11].

There was a relatively close fit between educational opportunity as expressed by the schools during the rapid expansion of elementary education in the second half of the nineteenth century and into the early years of the twentieth. But the elements summarized by Coleman were badly strained by the events that took place between World Wars I and II, especially industrialization, urbanization, population growth, and resulting increased expectations for secondary schools. Less and less were children and youth of diverse backgrounds attending the same school. With disparities in economic support of schools among communities, states, and regions becoming more and more apparent, it became increasingly difficult to defend equality of educational opportunity on a local basis.

What became strained immediately following World War I was the concept of a common curriculum at the secondary level. The high schools were called upon to serve a new clientele for whom the academic, college-preparatory curriculum appeared not to be appropriate. Only a small fraction of this new student body would go on to college. How to differentiate the curriculum for college-bound and non-college-bound youth became of nagging concern.

At the heart of the problem lay the issue of who could, would, or should go on to college. It is one thing to predict that 75 percent will not; another to predict specifically *who* will not; and quite another to reduce their options by preselection of a curricular track. Small wonder that the testing field blossomed.*

By World War II, the principle of comprehensive secondary education for all was well established. But the gap between the principle and reality was formidable. The fact that most youth attending high schools would not go to college had relatively little impact on the general conduct of secondary education. But, in the eyes of many, the academic curriculum was so watered down as to be inadequate preparation for college.

---

*It is of interest to note that the emergence of curriculum as a field of study also was vastly accelerated by the dilemmas of secondary education.

The problems and issues are still with us, in search of solutions. Meanwhile, developing countries around the world, on the brink of expanding secondary education rapidly, must make momentous decisions regarding vocational/technical schools (which cost several times more to construct, equip, and maintain), academic schools, or truly comprehensive schools. The road to enhanced educational opportunity appears to be that of transcending the schools to provide work experiences, apprenticeships, internships, and a variety of collaborations with commerce, government, industry, and the arts.

The flight to the suburbs, following World War II, sharply reduced the opportunity for children of varied backgrounds to attend the same school. One could still argue, defensively, that educational opportunity and, ultimately, access to a higher socioeconomic level were open; the outward trek simply was delayed for some. But inaccessibility to the local schoolhouse on the basis of color could not be denied. The concept of "separate but equal" was ruled unequal in 1954 by the Supreme Court of the United States.

Of great significance for any current analysis of educational opportunity, the Justices considered the psychological and educational *effects* of segregated schooling on the child. How is the concept of "equal" to be reconciled with "separate" if segregation itself in some way limits the benefits education is supposed to provide? This query has entered consistently into subsequent court cases regarding *de facto* segregation in the public schools.*

Coleman (1968) sharpens the implications of a concept of opportunity based on effects in an internal memorandum basic to the design of his monumental survey of educational opportunity (Coleman *et al.*, 1966). The following sentences from that memorandum are especially relevant: ". . . equality of educational opportunity is equality of results given *different* individual inputs. . . . Such a definition taken in the extreme would imply that educational equality is reached only when the results of schooling (achievement and attitudes) are the same for racial and religious minorities as for the dominant group" (1968, pp. 16-17). Subsequently, in the body of the report itself, Coleman and his associates discuss possible relationships between the population mix as a differential input in the educational setting and educational effects.

---

*For example, Crawford vs. Board of Education of the City of Los Angeles (peremptory writ of mandate, Superior Court of the State of California for the county of Los Angeles, February 11, 1970).

We have opened Pandora's box. It is one thing to view educational opportunity from the perspective of traditional inputs such as length of schooling, preparation of teachers, class size, funds spent, and the like. But the prospect of matching these and more subtle input variables to predictions of effects in planning and conducting educational programs, for individuals as well as groups, requires the elevation of educational practice to an applied science. Such a prospect provides direction for educational research of a kind the field has not enjoyed. Because of the importance and urgency of the issues, however, action is not likely to await research. The needed resources are more likely to go into a good deal of trial-and-error activity, rather than into the research efforts that might well be mounted at this time.

In the years since the 1954 Supreme Court ruling, socioeconomic class had become increasingly central to any considerations of educational opportunity and is destined to become even more so. We have come around full cycle: instead of schooling being the circuit-breaker, it has become part of the system for maintaining socioeconomic differentials. Although minority groups make up a vastly disproportionate part of our most economically disadvantaged, poverty is not confined within racial or ethnic boundaries. The costs of schooling relative to the property tax base reveal part of the problem. Quoting Berke (1971): "Variations in property tax base per pupil are immense. Ratios of four or five to one among areas in the amount of property per pupil are not at all unusual. The local property tax, therefore, makes it four or five times easier for some districts to raise a given amount of money from their own resources than it is for others [p. 13]."

But the cost of schooling goes up, regardless of disparities in ability to pay. James and Kelly (1966) point out that between 1930 and 1960 the national average cost per pupil rose 331 percent, but the per capita value of taxable property in large cities rose only 97 percent. Further, it costs more to provide less in the cities. Recognizing these problems, the Supreme Court of California in 1971 ruled on the illegality of the current system of property taxation in financing the public schools. This case and the United States Supreme Court case of 1954 together tell a significant story about the changing concept of educational opportunity in this country.

We have come a long way since Horace Mann. One is inclined to wonder where we would be today had we followed the path proposed by the Workingmen's Party of New York rather than that of the several fathers of the common school. In the late 1820s and early 1830s, the Party called for state-financed and controlled compulsory boarding schools and urged that each child be sent to such schools *as a means of creating a common environment* so that the education of all would be, indeed, equal (Lipset, 1970, p. 27). Today, of course, we would be required to add genetic manipulation to assure equality of readiness to profit from that environment and, following this, carefully controlled manipulation of the environment to assure equality of effects. With both kinds of interaction based on precise sciences and impeccably applied, will we have attained the ultimate in educational opportunity? And will we be beyond freedom and dignity? At any rate, I think we will be a little beyond what Horace Mann had in mind.

Educational opportunity in and through the schools

The preceding has drawn primarily upon the context within which schooling has proceeded, in sketching our changing perceptions of educational opportunity. It has suggested just a few of the implications for the conduct of schooling. Meanwhile, the schools have gone on about what they thought was their business.

In what follows, I have eschewed references both because of the vast literature on each of what might be termed a commonplace of schooling and because of the difficulty in selecting just a very few definitive but comprehensive items on each one. Because of space limitations, I have chosen only some of the most readily recognizable elements of schooling and a sparse array of examples.

The schools are and have been viewed in the United States as the major vehicle for educational opportunity. The rhetoric accompanying the drive for more and better schools has emphasized two educational aims, whatever else may have been stressed at various times: (a) preparation to participate constructively in the social order (to be able to read and understand the laws of the land, contribute to the G.N.P., and so on), and (b) development of individual potential for its own sake.

How well have the schools done in the light of these aims? In seeking to respond here to such a question, we are not concerned with whether these aims are appropriate or feasible for schooling. These are quite separate questions. We have grown up with the twin expectations, reiterated in every possible way, that the functions of the schools are to serve society and the individual. We are concerned, then, with the schooling gap—the gap between perceived expectations and perceived reality. Obviously, analysis of this gap brings into play personal interpretations of both the meaning of these educational expectations and conditions of schooling at any given time. The values are implicit.

Admissions policies throughout the hierarchy of schooling raise serious questions about individual opportunity, from definitions focusing on accessibility of schooling to definitions pertaining to effects. When a child is born determines whether he will begin school in the first grade a few days before his sixth birthday or a few days before his seventh. If the early years are as vital to subsequent development as most of the specialists in early childhood education suggest, then this delay of one-sixth of the life span already transpired could be of serious consequences. This limitation in educational opportunity could be corrected by establishing a birthday, let us say the fifth, nationwide as the date for legal access to school. Variations because of date of birth or place of abode, which have little or nothing to do with ability to take advantage of schooling, would be wiped out.

There is a vast literature on the abuses in *testing practices* and their impact on access to educational opportunity and, subsequently, the the amenities of life. Interest during the 1950s, on the use of tests in selecting children for limited places in public-school kindergartens illustrates an interesting quick turnabout in prevailing conceptions of educational opportunity. High test scores were used by some school psychologists as a basis for selecting those children thought to be *most ready* and *most able* to profit from earlier attendance at school. In the light of current thinking about disadvantage, it now would be more appropriate to use high test scores as a basis for *excluding* children [illegible] a limited number of places in kindergarten.

Before [illegible] the subject of testing, it is significant to point out that tests have been used disproportionately little as a means of providing alternative, more appropriate, learning opportunities for individuals. Tests have been part of the machinery for keeping the gates, for terminating passage through common learnings, rather than as windows on new possibilities. Only rarely have we used tests at the outset of a program of studies to excuse students from what they already know in order to pursue what they do not know. And so, those who do badly on tests frequently are barred from what they might profitably learn; those who do well often are provided a second opportunity to study what they already know. There is something so diabolically perverse about such a scheme that I seem to be denied the language for describing it.

It is difficult to reconcile very much of school *marking systems* with the dual aims of education cited earlier. One does not need to study Skinner's experiments to know the value of positive reinforcement in getting returns from instructional investments. And yet, 75 percent of the failing marks go to 25 percent of the pupils, most of whom are being negatively reinforced merely by attending school. And we can take little comfort from the supposed objectivity of the marking system. We know that in the humanities and the social sciences teachers' marks vary widely in judging the same pupil products. But even in mathematics, student marks frequently bear a disturbing correlation with the congruence or disparity between pupils' values and teachers' values.

School marks enter heavily and sometimes strangely into decisions of the socioeconomic marketplace. The boy who drops out of school carrying poor grades not only suffers the expected limitations in job and income but also, many places, must pay more for the insurance he needs for the car he drives to work. Furthermore, such a boy who needs a job and applies for a vacancy on the evening shift will have less chance of getting it than the boy with good grades who has not dropped out of school and does not need the job. Society, too often, reflects on the outside the school's abuses on the inside. These examples and many more become particularly disturbing when one reflects on the observation so effectively articulated by my colleague, C. Robert Pace, that grades seem only to predict grades, nothing else; not good work habits, not success in marriage, not dependability, not vocational success, not compassion.

Grouping practices have denied students access to stimulating peers, alternative learning opportunities, and upward mobility in the system, thus denying them some of the satisfaction they might have had while in school as well as some of the beneficial effects assumed to result from school. At the elementary level, children frequently are segregated into separate classes on the basis of some criterion of achievement, most often in reading. This practice erroneously is termed "ability grouping." The criterion used, however, is not often one of specific ability—a prediction of how the children are likely to perform—but rather a measure of how they already have performed. This is, indeed, a prediction but it is also a self-fulfilling prophecy. Although the results of research into pupil grouping are equivocal, one body of evidence stands up quite consistently: the children in the assumed-to-be-slow group do less well than their counterparts in heterogeneous classes.

Such practices often are used to give "academic" justification for racial segregation in presumably integrated schools. Consequently, it is not uncommon to find in predominantly *white* urbanized schools outside of the inner city "ability grouped" classes comprised almost entirely of *black* children. The school, too often, too faithfully reflects on the inside society's abuses on the outside.

It is necessary to distinguish sharply between grouping designed to bring students together for periods of time for achievement of specific purposes and grouping that segregates students for indefinite periods on a single criterion of assumed or apparent likeness. The former is useful and justifiable. One of the most serious and least examined limitations of the latter is that it deprives all students of the diversity in human background and personality to be derived from a truly "common" school. Further, following Coleman's analysis, such practices may very well hamper some students with respect to the academic goals of schooling. It appears, given certain extant knowledge and concepts of educational opportunity, that careful planning to assure heterogeneity rather than homogeneity in the internal structure and conduct of schooling is the more defensible policy.

The graded, lock-step structure of schooling and the *promotion practices* accompanying it play their part, too, in limiting educational opportunity. Contrary to the practices and supporting arguments of many teachers and principals, nonpromotion is not effective in enhancing pupil progress. Further, nonpromotion is not effected on the basis of rigorous criteria and hard data: whether or not a child is promoted depends heavily on the school he happens to attend and the teachers he happens to get. And, in the elementary school, three out of five nonpromoted pupils are boys; in the first grade, the ratio of boys to girls is even higher.

Some research suggests that the elimination of grades and the substitution of nongraded plans of school organization better reflect our knowledge of individual differences and facilitate conditions conducive to maximum learning. But efforts to change are hampered by tests and textbooks geared to grades and a host of expectations and practices that have cemented in the graded school. Educational practices devised for different times continue, in spite of the evidence against them, thwarting efforts to change toward schools that would reflect more faithfully changing knowledge and beliefs.

The relationship of the *curriculum* to educational opportunity is a topic so vast that it is foolhardy to pursue it here. Parents are very much preoccupied with their children's early successes in walking and talking which they frequently see as related to achieving parental goals. They pay relatively little attention to sensitivity to sound, color, and the aesthetic which are seen as at the fringes of parental goals. A pyramiding kind of reinforcement takes place, cutting off potential talents and shaping the others toward the apex. The elementary school picks up the process at a higher, narrower place on the pyramid. Science and social studies in the elementary school curriculum are very little of either; rather, they are the language arts with somewhat different content. And art is very little of art; it is a way of depicting social studies *cum* language arts in pictures.

Perhaps the most damning indictment of the curriculum, eloquently and repetitiously set forth by the neo-humanists, is that it fails to grip the student in vital, meaningful ways. Here, the expectations of many citizens have changed markedly from a future orientation to a "now" orientation. If the second of the dual aims cited earlier is to be interpreted as providing for self-selected student interests and goals, with accompanying rather than, at best, delayed gratification, then the school must be viewed as deficient.

Conclusion

Two sets of changing conditions have thrown formidable problems before our unrelenting drive for educational opportunity. The first is unprecedented technological and industrial growth with all the accompanying factors of urban crowding, mobility, and social disarray. These conditions have surrounded the schools and infused their operation with *de facto* segregation, gross inequalities in ability to support schools, and marked inequalities in what must be coped with and often circumvented if quality education is to be provided.

The second set is our expectations for educational opportunity. Have they become more enlightened? Or has the first set of conditions merely brought our shortcomings into sharper perspective? There probably is some truth in both implications; the relationships are exceedingly difficult to unravel. Clearly, neither meaningful analysis of educational opportunity nor constructive progress in providing it are possible if both are confined to educational systems and schools. It is a grave error, therefore, to condemn the schools alone for our present perceived shortcomings with respect to the provisions of educational opportunity and equality therein.

But it is an equally grave error to absolve the schools because of our hindsight in social, political, and economic realms or because of our ineptness in social engineering. For example, we often have chosen education as the scapegoat when we should have looked elsewhere, too, and educational means when we should have chosen solutions of social engineering. Nonetheless, we simply must not ignore the fact that the schools have too faithfully reflected inadequacies and inequalities of the larger social order, created a good many of their own, and succeeded in injecting some of their malfeasances into the surrounding society.

We have reached, with respect to educational opportunity, that dangerous and challenging time when conditions, concepts, and practices are in such disarray that virtually complete rethinking and reordering are called for. Tinkering and patching will continue; that is the nature of much human endeavor. But we need a fresh conceptualization and a new map. What do educational opportunity and equality of opportunity mean today? What should they mean? Where are we in the world of action with respect to this conception? Where should we be tomorrow, and what do we need to do to get there?

REFERENCES

Berke, J. S. The current crisis in school finance: inadequacy and inequity. *Phi Delta Kappan*, 1971, **53** (1), 2-7.

Coleman, J. S. The concept of equality of educational opportunity. *Harvard Educational Review*, 1968, **38** (1), 7-22.

Coleman, J. S., et al. *Equality of educational opportunity*. Washington, D. C.: U. S. Government Printing Office, 1966.

James, H. T., Kelly, J. A., & Garms, W. I. *Determinants of educational expenditures in large cities of the United States*. Cooperative Research Project No. 2389. Stanford University, School of Educational, 1966.

Lipset, S. M. The ideology of local control. In C. A. Bowers, I. Housego, & D. Dyke (Eds.), *Educational and social policy: local control of education*. New York: Random House, 1970. Pp. 21-42.

The System-Maintenance Function of Social Development in the Schools*

ROBERT D. HESS
Stanford University

In a time of social conflict and crisis, it is curious that the study of social development in the young is a neglected field in educational and psychological research. Perhaps this is because it has not the glamor of research on physiological bases of learning or the experimental elegance and relevance for academic achievement that are offered by research on cognition. Although there have been some studies of social processes as in the Robber's Cave experiment (Sherif, et al., 1961), development of a sense of moral justice (Kohlberg, 1963), and inquiries into self-concept (Coopersmith, 1967; Sears, 1964), there are few basic studies of children's concepts and behavior toward social issues, organizations, and institutions.

This situation could change very quickly, however, since scientific interest is occasionally responsive to obvious social crises. For example, the growth of research on cognitive development in poor and minority children in the past 10 years was stimulated by the public awareness of educational disadvantage of children from ethnic ghettos. If growth of research again follows evident need, social development may soon be a popular topic, for there is little doubt that this country is in trouble from racial discrimination and conflict, pollution and urban decay, poverty, and from political disaffection and dissent over the war. Perhaps there will soon be a comparable emphasis placed upon the relationship of social experience and behavior in the young to these stresses and defects in our institutions and systems. The purpose of this paper is to discuss some of the ways the public school is involved, both deliberately and inadvertently, in providing educational experience relevant to these social problems.

---

*I would like to acknowledge the contributions of the following people: Thomas Fox, who is working with me in the larger study; William McKenzie and Joseph McGeehan for their help in coding and analysis; and Lynn Curry, who summarized literature on development of social concepts. Paige Porter helped with revision of early drafts and conducted a separate study of the roles and images of women portrayed in social studies materials.

Social behavior traditionally refers to acts and feelings which relate persons as individuals to one another and to small groups—that is *person-relevant* behavior. There is another category of social activity, however, which might be called *system-relevant.* These are the attitudes and acts which relate individuals to the institution and to the social and political systems in which they live. This category is not easily described but it includes obvious political acts such as voting, protesting, helping candidates, as well as behavior which supports, changes, challenges, or erodes authority, tradition, and confidence in local institutions—the family, the school, the church, the university, the police and other agencies of government.

Such behavior is clearly of significance to the individual in terms of its immediate consequences to him, but its greatest importance comes from its function in maintaining or changing social organizations and national political systems. From this perspective, it appears that the current emphasis on humanistic experience, encounters, and inner feelings has not been paralleled with an emphasis upon the care and feeding of institutions and of social and political systems. Rather, social organizations have been blamed for many of our problems and attacked as archaic or evil. It is perhaps time to consider how they might be nurtured, and helped back to health if that is possible, and made more useful. This discussion, then, is oriented toward the development of social behavior in the young which has long term relevance for the vigor and effectiveness of the socio-political structures in which we live.

Research on system-relevant behavior has emerged in the past decade, primarily through studies of political socialization. Most of these studies have been conducted by political scientists and sociologists and have been concentrated upon political learning and the impact of socializing agents such as the family and the school upon emergence of attitudes toward political objects. Not only have our studies of system-relevant behavior been concerned with *normative* data, but also the findings have, for the most part, been interpreted in sociological and political science terms.

Although this recent research as well as the citizenship component of the National Assessment of Educational Progress (Education Commission of the States, 1970) provide useful normative information about attitudes toward political topics, we know relatively little about children's attitudes toward social problems and defects of the society. Also, many of the studies of political socialization were conducted before the counter culture, with its sensitivity to social problems and criticism of the establishment, became a significant movement in the United States. What seems to be needed now to make this growing field more useful to the schools is knowledge about the developmental processes involved in acquisition of social behavior, which affects the health of social structures. It is particularly critical to learn more about the age at which concepts about social issues are acquired and the sequence in which they develop and to determine to what extent conceptual learning and teaching have an impact upon development of behavior which will eventually maintain, change, or destroy social institutions.

It may be useful to illustrate the complexity of system-maintenance behavior by discussing what the schools are now doing and how the present practice reveals a need to accumulate a more useful knowledge base for educational planning.

In the United States as well as in other countries (Weiler, 1968; Hess & Torney, 1967; Bronfenbrenner, 1970; Ridley, Godwin & Doolin, 1971; Counts & Lodge, 1947), schools actively induct the young child into system-maintaining attitudes and behavior. Initially, these are of a type to produce feelings of attachment and nationalism, positive affect and regard for the school, the state, the nation and its officials. This is done both ritualistically and in incidental and subtle ways. The diffuse socialization input to which the young are exposed may include the salute to the flag and recitation of the pledge of allegiance; pictures of Washington, Lincoln, and other national heroes on the walls; observance of special holidays to honor national leaders and military victories; and the like. Later, however, the instruction becomes more deliberate, factual, and conceptual as it is incorporated into courses on history, civics, and government.

The intensity of national concern about such matters is indicated by the 1940 Supreme Court decision that students could be required to salute the flag, a decision reversed in 1943 (Oregon State Bar, 1968), firing of teachers thought to hold subversive ideologies, and community opposition to in-school activities thought to be openly opposed to our present form of government. Thus, the schools have been given a responsibility for shaping and producing system-maintaining behavior in children.

Some of the issues inherent in teaching system-relevant materials appear in the way some contemporary social studies texts deal with socio-political education and indoctrination. I would like to report some data on this point. Last spring, a graduate student, Tom Fox, and I began a study of the images that textbooks present of certain problems in American society.

We selected four areas in which stress or defects in the United States are evident and examined texts to see how they were described and presented. The four categories of potentially stress-producing social interaction are defined as follows:

1. *Race and ethnic relations* involves the social relationships that exist between the majority white population and groups that are distinguished by their culture or by their race. An example would be a discussion of the problems American Indians might encounter when moving from a reservation to a large city.

2. *Distribution of income, goods, and services* involves the dissemination of and access to the wealth and resources within our society. An example might be a discussion of the income and living conditions of migrant farm workers in this country.

3. *Political negotiations and processes* involves the interaction between groups of people, institutions, public officials, and decisions which involve allocation of resources or selection of political leaders. An example is a discussion of the tactics students or minority groups have used to secure certain political rights and benefits.

4. *Ecological practices* involves the utilization of resources and the preservation of the natural pattern of relations between man and his environment. An example is a discussion of problems created when chemicals and sewage are dumped into waterways.

We identified 18 textbooks adopted for 1970-1971 by at least two states on a statewide basis. Eight states were included—Alabama, California, Indiana, Mississippi, Nevada, Oklahoma, Texas, and West Virginia. In this study, which is the first phase of a larger analysis now underway in which books for fifth and ninth grades will also be examined, we read and categorized 17,286 paragraphs, which were our basic unit of analysis.

For each paragraph, we asked these questions: First, are any of these four problems mentioned in any way? The tally for this was a simple yes/no. A second question was applied to those paragraphs in which the category was introduced: Is this topic presented in a way that indicates there is a conflict or problem involved? If a Japanese gardener or art shopkeeper was mentioned with reference only to his business success as a member of an ethnic group but with no references to discrimination, this was scored as a no-conflict mention. Third, for paragraphs that included allusion to conflict or to the existence of a problem—poverty, dissent, pollution, etc.—the next evaluation was whether the author of the text presented the conflict or problem as one that was severe or mild.

The final evaluation was in terms of the degree to which the conflict or problem was presented as resolved or resolvable, as described by the authors of the texts. If there was, for example, an incident of exploitation of workers by an industry and the complaint was heard and a settlement agreeable to both worked out, the issue was resolved. A four-point scale was used to indicate the extent to which the reader was given the impression that the normal processes of the system were capable of resolving the difficulties described in the paragraph.

We thus had a *frequency count* of mentions for each category; a frequency of events described as involving conflict or difficulty; an intensity index of the severity of the problem; and a resolvability index, reflecting the impression of the system's adequacy for handling and resolving stress. The overall results are shown in Table 1.

A summary of this simple counting and rating approach shows that the appearance of social conflict in these texts is relatively low. Roughly speaking, only 17 percent of the 17 thousand paragraphs referred to these topics in some way, an average of about four percent per category. There was some variation among the four topics—economic and conservation topics were mentioned most often; political negotiation and processes least often. Although on the face of it, this frequency seems low, it is one of the interesting features of our state of theory and knowledge that we have developed virtually no criteria for judging whether these topics have received adequate coverage.

However that may be, more clearly significant is the way these topics were handled when they did appear. Of all paragraphs which dealt with these four sensitive issues, only 25 percent included any suggestion that conflict and stress might be involved, and in some categories allusions to conflict were even less frequent (Table 1).

Table 1

Number of Units Devoted to Four Social Issues
(Percentages)

Categories of Social Interaction	*Reference to Social Issues*	*Reference to Related Conflict*
Race and Ethnic Relations	3.93 (680)[a]	0.56 (97)
Distribution of Income, Goods and Services	7.18 (1241)	0.54 (93)
Political Negotiations and Processes	1.82 (315)	0.89 (154)
Ecological Practices	5.25 (908)	2.58 (447)
Total mentions	18.18 (3144)	4.27 (791)

[a] = n's in parentheses.
Total books coded = 18.
Total units (paragraphs) coded = 17,286.

Moreover, the frequency of mention of conflict appears to be inversely related to the intensity of the problem. Over half of the references dealing with conservation practices involved a suggestion of conflict or difficulty but only 12 percent of the references to race relations gave an impression that problems needed to be solved. To put it in more dramatic form, less than one percent of the 17 thousand paragraphs gave the reader the impression that race and ethnic relations in this country involve conflict and stress. Less than one percent discussed distribution of income in any way that suggests economic inequality or exploitation. There are references to credit buying, savings accounts, pay raises and the like, but few to poverty, unemployment, welfare, or hunger.

This general impression that we live in a benign society is extended in other ways. Not only are there few problems; those that do exist are not severe. The majority of the examples of conflict in the books we analyzed, regardless of which issue was involved—tended to be presented either as neutral (i.e., descriptive terms: 24 percent), or as being carried out in a spirit of cooperation (70 percent) as shown in

Table 2. Examples are "Groups will often campaign against a politician that they disagree with," or "People disagree about school integration." There is little hint of the sort of conflict that leads to bus burning, school boycotts, and fist fights among school board members. The image of polite exchange that is offered is more suited to an English tennis club than to the contemporary American scene.

Whatever the intensity of conflict, disagreement was usually handled in a fashion that led to resolution of the problems. On our

Table 2

Intensity of Conflict and Degree of Resolution in Reference to Four Social Issues (Percentages of mentions on all four issues)

Intensity and Direction of Conflict			*Degree of Resolution*		
(Assigned Values)					
Cooperation					
1.	(+3)	56.64 (448)[a]	1.	Resolved	48.30 (382)
2.	(+2)	9.99 (79)	2.	Near Resolution	44.37 (351)
3.	(+1)	4.30 (34)	3.	Far from Resolution	5.81 (46)
4.	(0)	23.64 (187)	4.	Not Resolvable	(0)
5.	(−1)	0.63 (5)	5.	Not Discussed	1.52 (12)
6.	(−2)	4.55 (36)			
7.	(−3)	0.25 (2)			
Hostility					
Total mentions with reference to conflict		100.00 (791)			100.00 (791)

[a] = n's in parentheses.
Total books coded = 18.
Total units (paragraphs) coded = 17,286.

four-point scale from "Resolved" to "Not Resolvable," 92 percent were rated as "Resolved" or "Near Resolution." Not one instance was described as not resolvable.

In summary, these major issues and sources of national tension were rarely presented as occasions of conflict or difficulty, were seen as relatively mild in intensity, and were all capable of resolution by the processes of our social and political system. A benign image of our society is being offered to nine and ten year old children. This type of socialization is oriented toward maintaining the status quo; it refuses to recognize defects, much less attempting to deal with them.[1]

Such a research summary is obviously only a partial statement of the problem. It is easy to point an accusing research finger at a possible flaw in the curriculum and textbooks; it is much more difficult to find a theory or evidence which would give reason to argue that they should be changed. Why should not the initial impression that a child gets of his country be that it is benign and that problems can be solved? What are the consequences of a more complete disclosure of our malignancies to the child at an early age? At this time, we do not know.

The task of socialization is relatively simple in a society where the purpose is to persuade the young to accept the values and patterns and behavior of adults. In a society which has severe conflict and stress, however, this simplistic process is no longer feasible as an educational policy. To maintain the status quo will also be to maintain and perhaps accentuate the internal stresses. What is needed, perhaps, is much more difficult—to help the young develop attachment to the institutions and political systems of the society and yet be able to change them sufficiently to deal with the sources of severe conflict.

---

[1] This general picture of social studies materials is consistent with other research findings, and there has been some pressure from political scientists and specialists in social studies curriculum for more realistic portrayals of both historic and contemporary problems of this country (*Harvard Educational Review*, 1968; American Political Science Association, 1971). The principal arguments behind this viewpoint are first, that the young become disillusioned when they discover that all is not as benign and ideal as the textbooks suggest and second, that present teaching often does not give the young citizen effective skills to deal with social problems. There is an additional argument with respect to the history of ethnic groups which deals with the nationalism, prejudice, and discrimination written into many representations of ethnic history and the treatment ethnic groups have received in the United States. This is a related problem, but the rationale is different and carries its own justification.

Such an orientation requires at least two steps. One is to evaluate our present techniques for developing social behavior in the young, and the second is to develop a knowledge base to assist curricular planning. At the present time, our information about the development of system-sustaining behavior in children is grossly incomplete. Indeed, the major question is whether the training and shaping of a young citizen has any effect upon the subsequent viability of a government and society. We have some heuristic theories about the function of socialization as a way to perpetuate existing systems but few about how the young can be taught to maintain and yet change a complex social order. Some argue that this cannot be done; that the power and pervasiveness of allegiance and nationalism make it impossible to prepare the young to change the system in effective ways. This argument holds that to encourage the young to believe in the system and the idea that it is responsive to them and to pressures for positive adaptive change is to mislead them. To change institutions is a formidable, frustrating, and long-term task; the eager young citizen may not only despair but, in profound frustration, drop out or turn his energies to destroying the system which he believes to be unresponsive.

A more specifically educational problem, however, has to do with the nature of the learning and teaching processes that are involved. The making of an effective citizen may need to utilize not only conceptual processes but conditioning and other forms of behavior shaping.

On the basis of present knowledge, for example, children apparently become attached to social and political units long before they have a conceptual grasp of either the process or the object of identification. Regard for the flag, respect for their country, response to the pledge of allegiance appear before clear concepts of government are present. That is, the process of national affiliation or attachment is essentially noncognitive or at least nonconceptual, in the usual sense of the term. It is an affective, gut reaction, not mediated by much information, by understanding, by informed choice or other processes in which the school specializes.

Similarly, compliance with rules and laws appears to be learned in significant part through conditioning, modeling, and other nonconceptual mediating operations. In short, two of the most basic social operations with regard to system-relevant behavior—national loyalty and compliance with authority—are apparently acquired in extraconceptual ways.

If attachment to the nation and compliance to law are produced by conditioning, identification, and modeling, rather than by conceptual learning and teaching, the school's function is deliberately manipulative as well as nonconceptual. This determination of the curriculum by the school obviously runs counter to current pressures to give children, even very young children, opportunities for self direction in selecting their own educational experiences.

The capability to comprehend the need for change, to make decisions and take effective action, however, requires a more conceptual grasp of problems and an understanding of the different consequences of alternative solutions. In short, the need to be critical and to examine the system and its faults calls upon different teaching and learning processes from those of earlier stages.[2]

There is some reason to believe, however, that most children are not able to deal with complex social concepts (the need for free speech, free press and religion, the rights of minorities, the logic of majority rule, etc.) until about the sixth grade or roughly the age of Piaget's formal operations. If additional research supports preliminary findings on this point it suggests that students will be approaching adolescence before they can handle the necessary concepts.

It is on these theoretical problems of cognitive readiness, sequence, and trainability that much research is now needed. It may be for example, that currently popular attempts to teach social science concepts in the early grades are not an effective use of time and may even interfere with the development of effective citizens. What seems to be clear is that our present procedures which are based largely on intuition and tradition seem not to be doing an effective job.

These are some of the problems that are encountered when we try to examine the empirical and theoretical base for social studies. The task with all its interdisciplinary facets is so formidable that only the threat to national survival or perhaps the protection of an ivory tower can give us courage to attack it.

---

[2] Even these are not necessarily restricted to conceptual experience. There is perhaps a need to experience discrimination to understand it fully, and to know what it means to be a member of a minority group in some situation in order to understand the social logic of the concept of minority rights. The processes, then, appear to include both affective and cerebral functions and experiences.

REFERENCES

American Political Science Association, Committee on Pre-Collegiate Education Report. *Political education in the public schools: the challenge for political science.* Washington, D. C.: American Political Science Association, 1971.

Bronfenbrenner, U. *Two worlds of childhood: U. S. and U. S. S. R.* New York: Russell Sage Foundation, 1970.

Coopersmith, S. *Antecedents of self-esteem.* San Francisco: W. H. Freeman & Co., 1967.

Counts, G. S., & Lodge, N. P. (translators). *I want to be like Stalin.* New York: John Day, 1947. [*Russian pedagogy* by B. T. Yesipov & N. K. Goncharov.]

Education Commission of the States. *National assessment of educational progress, citizenship: national results.* Denver: Educational Commission of the States, November, 1970.

Harvard Educational Review, **38**, No. 3, Summer 1968.

Hess, R. D., & Torney, J. V. *The development of political attitudes in children.* Chicago: Aldine Publishing Company, 1967.

Kohlberg, L. Moral development and identification. In H. W. Stevenson (Ed.), *Child psychology.* University of Chicago Press, 1963, Pp. 277–332.

Oregon State Bar. *The Flag Salute cases. Liberty and the law: case studies in the bill of rights.* Englewood Cliffs, N. J.: Prentice-Hall, 1968.

Ridley, C. P., Godwin, P. H. B., & Doolin, D. J. *The making of a model citizen in Communist China.* Stanford: Hoover Institution Press, 1971.

Sears, P. S. Self-concept in the service of educational goals. *California Journal for Instructional Improvement*, 1964, **7**, 3–17.

Sherif, M., et al. *Intergroup conflict and cooperation: the Robber's Cave experiments.* Norman, Okla.: University of Oklahoma, Institute for Group Relations, 1961.

Weiler, H. N. Education and political development. *Review of Educational Research*, 1968, **37**, 231–242.

Phyllis A. Wallace emphasizes the influence of "human capital" on educational opportunity and student achievement.

O. Meredith Wilson invokes his favorite philosopher in discussing Session I.

Economics and the Values of Society

PHYLLIS A. WALLACE
Metropolitan Applied Research Center, Inc.

Introduction

Education is perceived by most Americans as a major instrument for building a pluralistic society. Some of the desired educational outcomes, at least in the long run, are wider occupational choices, increased lifetime earnings, expanded social and political participation, and learning of common cultural values and moral choices. Until recently, economists have been almost single-mindedly concerned with those activities which enhance economic productivity. The effect of schooling on earnings has been substantiated in a number of studies (Ashenfelter & Mooney, 1968; Hansen, 1963; Hansen & Weisbrod, 1970; Michaelson, 1968; Schultz, 1963; Weiss, 1970). Beginning in the latter part of the 1960s, however, a few economists attempted to formulate theories about and to undertake empirical investigations on the complexities of providing education to lower status children, especially those in urban school systems.

The growing inability of urban primary and secondary educational systems to perform effectively for minority and poor children can be examined from several economic perspectives—the inequality of economic resources, the investment of human capital, and education production functions. My brief survey of findings from economic research substantiates much of what may already be accepted by psychologists. It is useful to review these studies in detail before I comment on the two questions: (a) Can we afford an educational program designed to provide increasing opportunity in the future? and (b) What will happen if we do not provide such an opportunity?

Inequality of Economic Resources

Findings from several studies on the marked inequality of economic resources available to schools serving lower status children were presented in statements before the Senate Select Committee on Equal Educational Opportunity (the Mondale Committee) (*Equal Educational Opportunity*, 1970; Guthrie et al., 1970; Levin, 1970b). These studies have been published in Part Seven of those hearings (*Inequality of Economic Resources*). The disproportionate flow of local, state, and federal funds to schools servicing middle class families means that poor children have fewer dollars spent on them. These fiscal disparities are translated into disparities in service. Thus, the quality of school services received is determined by the socioeconomic status of the child.

Levin and other researchers have shown that lower quality school services are associated with lower levels of achievement (student performance on tests of cognitive ability), and vice versa (Guthrie et al., 1970; Levin, 1970b). Although academic achievement may also be influenced by conditions external to the school, the quality of the school remains a primary determinant of pupil achievement. The recent decision of the Supreme Court of the State of California (1971) addressed itself mainly to the effects of differential school revenues derived from local property taxes. The Court stated:

> . . . so long as the assessed valuation within a district's boundaries is a major determinant of how much it can spend for its schools, only a district with a large tax base will be truly able to decide how much it really cares about education. The poor district cannot freely choose to tax itself into an excellence which its tax rolls cannot provide [p. 47].

Poor school districts, however, may be discriminated against even where state and federal equalization funds are available. There are a number of case studies of these compensatory funds being siphoned off to other programs or invested in outmoded remedies (NAACP, 1969). Although the prevailing system of financing primary and secondary education has produced an unequal investment in the education of poor children, a reduction in the resource inequities alone would not be sufficient for equalizing the academic achievement of these children. We need to know how to improve the efficiency of schools.

Phyllis A. Wallace

Investment in Human Capital

Beginning in the latter part of the 1950s, economists started to formulate theories on investment in human capital (Becker, 1964; Schultz, 1967; Weisbrod, 1962). Initially, they were concerned with the private and social costs and returns of investing in higher education and training. Increased individual lifetime earnings, wider occupational choices, more equitable distribution of income were some of the expected benefits. Even if academic achievement in the lower grades is considered only as a proxy for later economic performance, the relationships are not as strong for blacks as for whites.

Many lower status children begin school with major disabilities that economists define as a deficit in capital embodiment. Levin states that populations from lower social class backgrounds and poorer areas of the country generally have received poorer health services, poorer prenatal care for their children, poorer nutritional services and less educational investment from their parents and communities (*Equal Educational Opportunity*, 1970).

Guthrie has estimated that a child whose mother has not graduated from the eighth grade will have about $4,000 of educational capital embodied in him prior to going to school, and a child whose mother has graduated from college will have about $16,000 to $17,000 of home or family education embodied in him (*Equal Educational Opportunity*, 1970). The values of the parents' contributed educational endowment are excellent predictors of academic success at grade one. Differences in human capital embodiment explain approximately 95 percent of the variance in pupil verbal skills for white first graders and 88 percent of the variance for nonwhite first graders (Dugan, 1969). Even if identical amounts are spent on the in-school education of the children from the different backgrounds, a discrepancy of $20,000 to $30,000 might remain at the end of the twelfth grade.

The critical variable is seen as the social investment made by the school. To the extent that human capital investments by family and community in individuals from different social groupings varies significantly, the question can be raised about the capability of the schools to equalize opportunities.

Essentially, we wish to know what amount of additional school investment is needed to equalize student academic achievement for lower status children whose peers have received a heavy investment from their parents.

Quality of Schools

The statistical analysis by Bowles and Levin of the Coleman data strongly suggests that poor measurement of school resources, inadequate control for social background, and inappropriate statistical techniques tended to underestimate the importance of teacher and other school resources on student achievement (Bowles & Levin, 1968). Preliminary results from the Bowles study of black high school twelfth graders indicate that verbal ability of teachers is the most important explanatory variable for student achievement. Student attitudes toward self and toward control of environment were also powerfully related to achievement (Bowles, 1970). These estimates were derived from an educational production function and should be treated with caution. Measures of cost effectiveness (maximizing educational output for any budget constraint) are likewise rudimentary (Carpenter & Haggart, 1970; Hanushek, 1970; Levin, 1970a). Nevertheless, these efforts represent the required first steps to understanding some of the paradoxes of public education in America.

Economists and others have explored whether the assessment of student performance in the class can be used as a standard for evaluating school performance. However, neither rewards for attaining particular goals nor penalties for failing are built into the present system. Performance contracting, voucher financial systems, and decentralization or community control of schools are being tested to determine whether they can effectively improve the ability of the schools to respond to the needs of lower status children. There are major shortcomings in all of these approaches to accountability.

Performance or incentive fee contracting is designed to permit suppliers of educational services to try to raise achievement levels of children by a specified amount within a set time limit. The first results from the Office of Economic Opportunity's $6.5 million nationwide experiment in performance contracting were recently reported as costing private companies more than regular school systems to teach basic subjects to failing students. Although the private companies showed higher achievement gain than regular school programs, in most cases the students did not exceed predetermined achievement levels (New teaching technique, 1971). The social burden of educating these children is inflated by the fact that dual systems, school and performance contractors, are paid. The four year experiment in one school in Gary shows significant improvement in skills in reading and mathematics for the first year (King, 1971).

The voucher or market approach to providing education was suggested almost twenty years ago by the University of Chicago economist, Milton Friedman. The Friedman approach would provide the parents of the child with tuition vouchers equal to the present amount being spent; they could be applied to any school approved by the State. Friedman was interested in increasing competition in the educational marketplace. The purchase of educational services from private enterprises would not necessarily produce significant increases in educational effectiveness and output for lower status children. It is likely that social segregation would be increased and that present disparities in opportunities afforded to different social and racial groups would be increased (Levin, 1968). One way to overcome these handicaps would be to grant vouchers only to students from poor families.

It is too early to tell whether the decentralized school districts will be more successful in ensuring that quality education is available to ghetto school systems and that parental and community involvement will make the schools more accountable to the populations they serve (Levin, 1970c).

The Clark plan for the Washington, D. C., school system (*A Possible Reality*) presented a model for the establishment of systemwide accountability (Metropolitan Research Center, 1970). It assumed that managers of the system could and should modify their own behavior in ways that would be beneficial to the system. There are certain inherent risks of self-analysis and regulation.

The shift toward predictive profiles of performances based on estimates of likelihood of success given inadequacies of home and school environment, would establish low expectations. We should test school systems' effective responsibility in teaching students to read, not by measuring difficulty and challenge of their tasks, but by measuring the accomplishment or failure of students.

The emerging concerns about accountability also arose as taxpayers refused to raise local taxes to support higher school budgets. Thus, fiscal accountability to state legislatures rather than educational accountability to parents may be achieved first.

Although there is a strong relationship between both academic attainment and achievement and earnings, there is now a considerable body of evidence from the economic literature suggesting that economic returns to schooling are significantly less for black than for white students (Ashenfelter, 1966; Hanoch, 1967; Levin, Guthrie, Kleindorfer, & Stout, 1971; Mincer, 1970; Ribich, 1968; Weiss, 1970). Employment discrimination as measured by the ratio of expected lifetime earnings of whites with a given level of schooling to the expected lifetime earnings of blacks with a similar level of schooling shows that blacks benefit less than whites in terms of increased income. Returns from additional education are much lower for blacks than for whites when academic achievement is used in deriving the earnings function. In a study of black-white differences in returns to schooling, Welch concludes that more recent black graduates fare much better in comparison with whites than do older blacks. For urban males entering the labor force prior to 1960, relative black income falls with increased schooling, but for more recent entrants, relative black income increases with schooling (Welch, 1971). The detailed 1970 census data should throw more light on the ambiguous role of education for blacks as reflected in their enormous difficulties in labor markets.

In his analysis of the external benefits of public education Weisbrod (1964) states that "literacy has a social value beyond its value to the individual possessing it and to employers. . . . If equality of education is a social goal, then education pays social returns over and above the private returns to the recipients of the education, although there may be some private costs to previous holders of privileged positions [p. 33]." Economists and others frequently have restricted their concerns to the distribution of broad based tax burdens for the support of education as compared to the more limited distribution of educational benefits. It is clear that some significant social benefits of education may be associated with reduction of the costs of crime, welfare and discrimination of all kinds. These social benefits need to be identified and measured, even if it is difficult to impute market values.

Conclusion

The achievement of high quality education for all assumes single systems of goals for behavioral achievement, and single systems of tests and measures to determine the effectiveness of our schools. I am not certain that there are economic answers to the questions raised earlier in this report. Undereducation (education in the sense of perception of society and perception of self in that society) of a sizeable part of the population increases the alienation and isolation of this group from the total society.

It is appropriate to conclude with comments from two social scientists who have raised difficult questions about moral, aesthetic, and intellectual judgment.

In his 1968 presidential address before the American Economic Association, Kenneth Boulding (1970) stated:

> Even though economic measurement may be abused, its effect on the formation of moral judgments is great and on the whole, I believe, beneficial. For instance, the whole idea of cost-benefit analysis in terms of monetary units, say "real" dollars of constant purchasing power, is of enormous importance in the evaluation of social choices and even of social institutions. We can grant, of course, that the "real" dollar which oddly enough is a strictly imaginary one, is a dangerously imperfect measure of the quality of human life and human values. Nevertheless, it is a useful first approximation, and when it comes to evaluating difficult choices it is extremely useful to have a first approximation that we can modify. Without some guideline, indeed all evaluation is random selection by wild hunches. It is true, of course, that cost-benefit analyses, whether of water projects or other pork barrel items or, in more recent years, weapon systems, can be manipulated to suit the prejudices of people who are trying to influence the decisions. Nevertheless, the fundamental principle that we should count all costs, whether easily countable or not, and evaluate all rewards, however hard they are to evaluate, is one which emerges squarely out of economics and one which is at least a preliminary guideline in the formation of moral judgment in what might be called the "economic ethic" [pp. 129–130].

In the 1971 Sol Feinstone Lecture at Syracuse University, Kenneth B. Clark (1971), then President of the American Psychological Association, stated:

> Human beings without the minimal education essential for a constructive role in society are doomed to that profound frustration which demands direct or vicarious forms of hostility, aggression and cruelties . . . The intensified struggle to upgrade the quality of education provided for the children in the public schools of American urban ghettos and slums must be understood as part of this seemingly endless and agonizing struggle to provide human beings with the realistic base for genuine pride, self respect, and that affirmative freedom and human dignity consistent with personal choices and options. These are the stabilizing ingredients essential for a society that is democratic . . . [pp. 2, 3].

REFERENCES

Ashenfelter, O. Minority employment patterns. Prepared by U. S. Equal Opportunity Commission and the Manpower Administration of the Department of Labor, 1966.

Ashenfelter, O., & Mooney, J. Graduate education, ability, and earnings. *Review of Economics and Statistics*, 1968, **50** (1), 78–86.

Becker, G. S. *Human capital*. New York: Columbia University Press, 1964.

Boulding, K. E. Economics as a moral science. In *Economics as a science*. New York: McGraw-Hill, 1970. Pp. 117–138.

Bowles, S. Towards an educational production function. In L. W. Hansen (Ed.), *Education, income, and human capital*. New York: National Bureau of Economic Research and Columbia University Press, 1970. Pp. 11–61.

Bowles, S. J., & Levin, H. The determinants of scholastic achievement: an appraisal of some recent findings. *Journal of Human Resources*, 1968, **3** (1), 3–24.

Carpenter, M. B., & Haggart, S. A. Cost-effectiveness analysis for educational planning. *Educational Technology*, 1970, **10**, 26–30.

Clark, K. B. Education: the basis of freedom. The 1971 Sol Feinstone Lecture, Syracuse University.

Dugan, D. J. The impact of parental and educational investments upon student achievement. New York: American Statistical Association, 1969 (Mimeographed).

Equal educational opportunity. Hearings before Select Committee on Equal Educational Opportunity of the United States Senate, 91st Congress, 2nd Session, Part 7. Washington, D. C : U. S. Government Printing Office, 1970. Pp. 3401–3681.

Guthrie, J. W., et al. Educational inequality, school finance, and a plan for the 70's. Reprinted in *Equal educational opportunity*. Pp. 3451–3500.

Hanoch, G. Personal earnings and investment in schooling. *Journal of Human Resources*, 1967, **2**, 310–329.

Hansen, L. W. Total and private rates of return to investment in schooling. *Journal of Political Economy*, 1963, **71** (2), 128–140.

Hansen, L. W., & Weisbrod, B. Schooling and earnings of low achievers. *American Economic Review*, 1970, **60** (3), 409–418.

Hanushek, E. *The value of teachers in teaching*. RAND (RM 6362 CC/RC), December 1970.

King, S. S. Pupils improve in school run by private company. *New York Times*, September 29, 1971. Pp. 1, 42.

Levin, H. M. *The failure of the public schools and the free market remedy*. Washington, D. C.: The Brookings Institution, 1968. Reprint No. 148.

Levin, H. M. A cost-effective analysis of teacher selection. *Journal of Human Resources*, 1970, **5** (1), 24–33. (a)

Levin, H. M. Financing education for the urban disadvantaged. Reprinted in *Equal educational opportunity*. Pp. 3612–3621. (b)

Levin, H. M. (Ed.). *Community control of schools*. Washington, D. C.: The Brookings Institution, 1970. (c)

Levin, H. M., Guthrie, J. W., Kleindorfer, G. B., & Stout, R. School achievement and post school success: a review. *Review of Educational Research*, 1971, **41** (1), 1–16.

Metropolitan Applied Research Center. *A possible reality*. A design for the attainment of high academic achievement for the students of the public elementary and junior high schools of Washington, D. C., June 30, 1970.

Michaelson, S. Incomes of racial minorities. In *Studies of social economics*. Washington, D. C.: The Brookings Institution, 1968.

Mincer, J. The distribution of incomes: a survey with special reference to the human capital approach. *Journal of Economic Literature*, 1970, **8** (1), 1–26.

New teaching technique is costly. *Washington Evening Star*, September 20, 1971. Pp. 1, 6.

Ribich, T. *Education and poverty*. Washington, D. C.: The Brookings Institution, 1968.

Schultz, T. W. *The economic value of education.* New York: Columbia University Press, 1963.

Schultz, T. W. The rate of return in allocating investment resources to education. *Journal of Human Resources*, 1967, **2** (3), 293–309.

Supreme Court of the State of California. *John Serrano, Jr., vs. Ivy Baker Priest*, 1971. Pp. 47.

Title I of ESEA: is it helping poor children? New York: NAACP Legal Defense and Education Fund, 1969.

Weisbrod, B. Education and investment in human capital. *Journal of Political Economy*, 1962, **70** (5), Supplement, 106–123.

Weisbrod, B. *External benefits of public education: an economic analysis.* Princeton: Industrial Relations Section, Princeton University, 1964.

Weiss, R. D. The effect of education upon the earnings of whites and blacks. *Review of Economics and Statistics*, 1970, **52** (2), 150–159.

Welch, Finis. *Black-white differences in returns to schooling.* Presented at conference on discrimination in labor markets, Princeton University, October 7–8, 1971.

Discussion

O. Meredith Wilson
Center for Advanced Study in the Behavioral Sciences

This first session has been asked to provide the setting—educational, historical, social, and economic—within which testing must operate. There can be no great value in my summarizing what has already been presented with such clarity by three distinguished speakers. Perhaps, therefore, I may be forgiven if in my remarks I am content to expand, elaborate, or reinforce what has already been said.

Our historical view of equal educational opportunity was always flawed. Our society began at a time when class structure was an accepted reality and the typical English class structure which had been transplanted here was complicated by the intrusion of slavery as an integral part of our social life and productive force. The end of the Civil War did not bring with it a transformation in the minds of educational philosophers. Much of the conversation about equal educational opportunity could and did proceed as though the American black did not even need to be a part of the conversation or of the national concern. As a result, the history of educational opportunity looks very different if analyzed in the Mesabi Range of northern Minnesota from what it would appear if analyzed in the cotton country of the deep south. For the Finn, the Czech, the Serb, or other European coming to the Mesabi Range, the first job of education was unification. The chore of sorting and separating could only come after a community was built. The elementary grades did provide a common language for children whose parents still depended upon ethnic social clubs for their social life and continued to speak an eastern European language in the home. The parents did consider upward mobility as a reality to be achieved, if not by themselves, then at least by their children. Schools were recognized as a ladder on which those children could climb, and the quality of the schools was of great concern to them. Absentee mining companies did have substantial properties that could be assessed substantial local taxes, and palatial schools, such as the high school in Hibbing, Minnesota, are a monument to these peoples' faith in education and their political skill in achieving appropriations to satisfy their educational aspirations. One needs only to review the roster of state legislators in the State of Minnesota to realize that upward mobility for these pople did not mean economic prosperity alone. It meant turning the social classes on their heads so that the children of immigrant miners could be the mayors, or legislators, or governors, within the new society. The schools did provide a common core of language, of political and social tradition and respect, and did advance the students' views to the point where they could leave the mines for the skilled laboring force or to enter professions.

Perhaps the most graphic account of the upward mobility promised by this society came in the very first quarter century of our cultural experience and was recorded in the diary of John Winthrop (*The People Shall Judge*, 1949):

> (APRIL 1645)
>
> The wars in England kept servants from coming to us, so as those we had could not be hired, when their times were out, but upon unreasonable terms, and we found it very difficult to pay their wages to their content (for money was very scarce). I may upon this occasion report a passage between one of Rowley and his servant. The master, being forced to sell a pair of his oxen to pay his servant his wages, told his servant he could keep him no longer, not knowing how to pay him the next year. The servant answered, he would serve him for more of his cattle. But how shall I do (saith the master) when all my cattle are gone? The servant replied, you shall then serve me, and so you may have your cattle again [p. 14].

In the margin of his journal, opposite this paragraph, Governor Winthrop had written "insolent." For those poor who were to be allowed to aspire to upward mobility, the schools were the means by which what Winthrop had considered "insolent" could become a legitimate ambition, and a realistic vision of the future. Yet side by side with a system of education that would allow the white dispossessed to climb, our society condoned and maintained a system within which, for a black, striving to climb was still held to be "insolent."

On the Mesabi Range the most important lesson of democracy was taught, not through the curriculum, but through the social situation. The mayor, the mine superintendent, and the immigrant miners all sent their children to the same school, and from nine in the morning until three-thirty at night, the miner's child could learn the lesson of aspiration as he noted that he and his friends, the children of immigrant miners, were entitled to and received as much attention from the teacher, a public servant, as did the child of the school superintendent or the child of the local banker. Moreover, the son of the banker, the mining superintendent, or the schoolmaster learned another lesson of democracy—the lesson of humility, since the public school taught him that he had no more claim on the time of the teacher than did those who were still struggling to learn the language. Until the case of Brown versus the United States, no such opportunity to study in the same classroom and to learn the lessons of democracy was provided the American southern blacks. Many of our present problems in education and much of our concern with social justice are born from the fact that we wrote and read our histories of education as though education operated and affected the dominant class only. Because we deceived ourselves into believing we were a classless society, we affected the status of the Negroes even more adversely than had we openly reproached them. We visited upon them the ultimate insult of omitting them from consideration at all in our history or our policy.

Now there does seem to be some desire to make equality of education really mean equality of opportunity. Mr. Goodlad has raised the question whether we can hope for an equality of education as output and has called attention to an internal memorandum of the Coleman Committee which raised this issue. As Goodlad says, there seems very little hope that education will ever become so precise a science that it could take ever so differently qualified students into the system and grind them out as equals at the end of some twelve or sixteen years of instruction. But just raising the question does justify him in speculating whether high test scores should not henceforth be used to exclude students from kindergarten rather than to admit them.

Dr. Hess has posed some of the most serious questions I have ever seen placed before a conference, and they will require the attention of more than the testing fraternity. I found his questions peculiarly distressing because I received and read his paper very shortly after having reviewed B. F. Skinner's *Beyond Freedom and Dignity*. The evidence of the research of Hess and Fox points to the fact that our first years of education are clearly aimed at system maintenance. We have conceived the social studies curriculum for the elementary years as an appropriate way to sustain or sell our social institutions at first ceremonially and later substantively. What the meaning of freedom in subsequent years might be if the first five are carefully plotted and planned to condition the minds of children is a matter for each of us to judge. Clearly an education calculated to inculcate existing values must surely be designed to avert or to retard a change. Inculcation may be appropriate up to some age; but if the first few years of education are a period of mind management, one is tempted to ask with anxiety: how is the subsequent manipulation of my mind disguised? You remember that Mr. Skinner (1971) has said, "We recognize a person's dignity or worth when we give him credit for what he has done. The amount of credit we give is inversely proportional to the conspicuousness of the causes of his behavior [p. 58]."

I am sure Skinner would expect a man trained in history, as I am, to respond negatively, as I do, to his substitution of reinforcement, aversion and predictability behavior for the myth of freedom and dignity which so reinforces my spirit. And I cannot resist giving you a quick sample of the heroic which I cherish, and which I am loath to lose.

One of the most eloquent statements, and still brief enough to impose on you, of the autonomous man I find in Epictetus, Book 1. It is an exchange between Helvidius and the Emperor Vespasion:

Helvidius: It is in your power not to allow me to be a member of the Senate, but so long as I am I must go in.

Vespasion: Well go in then, but say nothing.

Helvidius: Do not ask my opinion and I will stay silent.

Vespasion: But I must ask your opinion.

Helvidius: And I must say what I think right.

Vespasion: But if you do I will put you to death.

Helvidius: When did I tell you I was immortal? You will do your part, and I will do mine; it is your part to kill; it is mine to die, but not in fear; yours to banish me; mine to depart without sorrow.

I know of course that I am in trouble because programmed learning uses the technique of testing to move people in an orderly fashion from ignorance to understanding, and it was programmed learning that first introduced me to the language of reinforcement and aversion, but I still stubbornly want my freedom, and to retain the catalogue of heroes of which Helvidius is a sample.

Dr. Phyllis Wallace could have begun her paper by saying that if being accepted and treated hospitably by the dominant society were a necessary reinforcement, and if reinforcing were important to education, then we have incurred a heavy debt in respect to the American black. We have alleged that there is a relationship between the amount of education an individual receives and his earning power in later life. Though that conviction is held by most of our society, we have nevertheless provided less education to the American black. Once he has finished his education, we have given him less opportunity for full employment than his brothers, and once we have employed him, we have paid him less for his hours of work than we do those of a different color who are no better prepared. To erase all these disadvantages, education alone is not enough but it is still of tremendous importance.

Dr. Wallace introduced an interesting concept which may be common among economists but is new to me. She said, citing Guthrie, "Differences in human capital embodiment explain approximately 95 percent of the variance in pupil verbal skills for white first graders and 88 percent of the variance for non-white first graders." (The human capital embodiment is a figure derived from the education of the parents of the child. For example, "a child whose mother has graduated from college will have $16,000 to $17,000 of home or family education embodied in him.")

The most important social agency available to overcome this handicap in human capital is the school, and apparently the most important factor within the school, if Dr. Wallace's information is correct, is the verbal skill of the teacher. If testing services have not already confirmed these judgments upon which Dr. Wallace depends, then one of the critical products that could be provided to us in education by the testing services is a clear assessment of their validity. Meanwhile it is particularly important to recognize that those who arrive at school with a deficient human capital embodiment are usually condemned to attend the least well financed schools. It is a fair generalization to say that where the poor live, the assessed property value is low, and it is virtually impossible to raise a substantial amount of local money to improve the schools. Yet it remains true that we leave the burdens of carrying the schools to the real estate and property taxes, assess them locally and distribute the money only within narrow school districts, so that the poverty of the people is reinforced by the poverty of the schools. At the same time, the federal government participates in only seven percent of the costs of public education. In historical, social, and economic perspective, it may be said that as the world has changed in its dependence on knowledge, we have remained too rigid in our adaptation to school administration and school finance.

The brightest hope that I have observed is found in the legal argument of the plaintiffs and the decision of the California Supreme Court in the case of John Serrano Jr. vs. Ivy Baker Priest (California Official Reports, 1971). It has been referred to by earlier papers. I should like to read part of the case into the record.

> The complaint alleges that the financing scheme:
>
> A. Makes the quality of education for school age children in California, including Plaintiff Children, a function of the wealth of the children's parents and neighbors, as measured by the tax base of the school district in which said children reside, and
>
> B. Makes the quality of education for school age children in California, including Plaintiff Children, a function of the geographical accident of the school district in which said children reside, and
>
> C. Fails to take account of any of the variety of educational needs of the several school districts (and of the children therein) of the State of California, and
>
> D. Provides students living in some school districts of the State with material advantages over students in other school districts in selecting and pursuing their educational goals, and
>
> E. Fails to provide children of substantially equal age, aptitude, motivation, and ability with substantially equal educational resources, and
>
> F. Perpetuates marked differences in the quality of educational services, equipment and other facilities which exist among the public school districts of the State as a result of the inequitable apportionment of State resources in the past years.
>
> G. The use of the "school district" as a unit for the differential allocation of educational funds bears no reasonable relation to the California legislative purpose of providing equal educational opportunity for all school children within the State.
>
> H. The part of the State financing scheme which permits each school district to retain and expend within that district all of the property tax collected within that district bears no reasonable relation to any educational objective or need.

> I. A disproportionate number of school children who are black children, children with Spanish surnames, children belonging to other minority groups reside in school districts in which a relatively inferior educational opportunity is provided [p. 590].
>
> For example, Baldwin Park citizens, who paid a school tax of $5.48 per $100 of assessed valuation in 1968-1969, were able to spend less than half as much on education as Beverly Hills residents, who were taxed only $2.38 per $100. (Cal. Dept. of Ed., *op. cit. supra.* Table III-16, p. 43) [p. 598].
>
> In summary, so long as the assessed valuation within a district's boundaries is a major determinant of how much it can spend for its schools, only a district with a large tax base will be truly able to decide how much it really cares about education. The poor district cannot freely choose to tax itself into an excellence which its tax rolls cannot provide. Far from being necessary to promote local fiscal choice, the present financing system actually deprives the less wealthy districts of that option [p. 611].
>
> Finally, education is so important that the state has made it compulsory—not only in the requirement of attendance but also by assignment to a particular district and school. Although a child of wealthy parents has the opportunity to attend a private school, this freedom is seldom available to the indigent. In this context, it has been suggested that "a child of the poor assigned willy-nilly to an inferior state school takes on the complexion of a prisoner, complete with a minimum sentence of 12 years (Coons, Clune & Sugarman, *supra*, 57 Cal. L. Rev. at p. 388) [p. 610]."

It is still true that if our society wishes to be open and to provide mobility for all of its people, the school is the most effective ladder on which the disadvantaged can climb; that is, it is if every man has such a ladder planted firmly on his own ground so that he can, if he will climb on it to pluck the highest fruits of our economy and our society.

REFERENCES

California official reports, No. 25. San Francisco: Bancroft-Whitney Co., 1971.

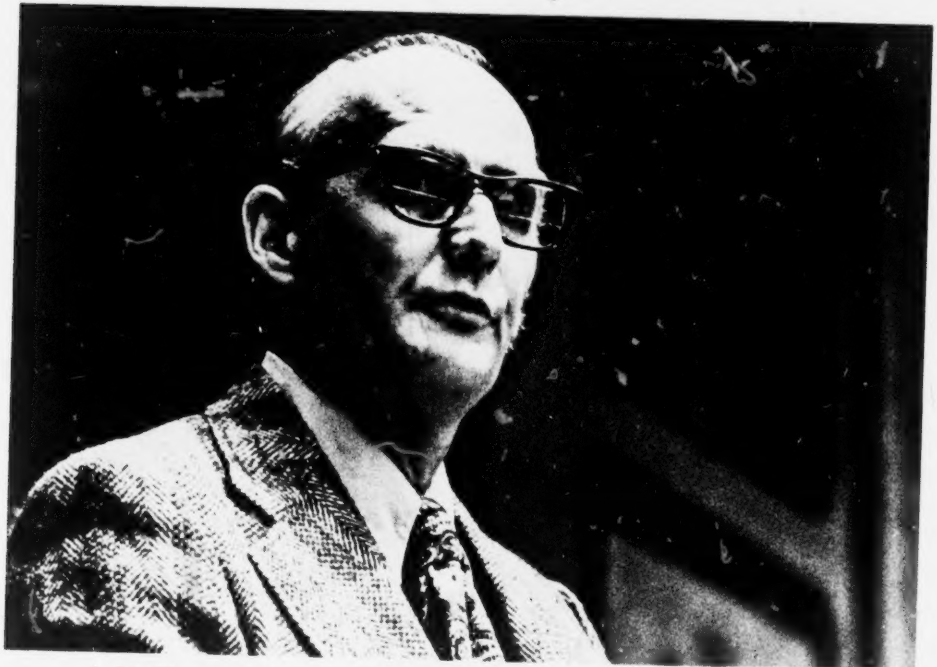
Skinner, B. F. *Beyond freedom and dignity*. New York: Alfred A. Knopf, 1971.

The people shall judge, Vol. 1. Chicago: University of Chicago Press, 1949.

Session II

The Measurement Context: Past, Present, and Future

Philip H. DuBois traces measurement's contributions to educational opportunity.



A. N. Hieronymus reviews the capabilities of present-day measurement.

Increase in Educational Opportunity Through Measurement

PHILIP H. DUBOIS
Washington University

When we examine the records left by early inventors of psychological and educational tests, we find that the motives they ascribed to themselves for their work varied enormously. Galton and Pearson developed tests as aids in their studies of heredity. James McKeen Cattell attempted to study the range and importance of individual differences in behavior. Joseph Mayer Rice developed his achievement tests in spelling, arithmetic, and language as tools in support of educational reform, while Charles Spearman constructed tests in his attempts to discover the nature of human mentality. The motives of E. L. Thorndike in inaugurating and executing a highly dynamic program in test development were undoubtedly complex, but it is apparent that a number of his students undertook test development for a relatively simple reason—to get material for a Ph.D. dissertation. Of the very early testers only Binet and Witmer seem to have developed a primary interest in using the results of tests in the diagnosis of individuals, but in the case of Binet this seems to have come after a long period of interest on testing centered around scientific purposes.

It is the thesis of this paper that, since their inception three quarters of a century ago, modern educational tests with their special features of comparability and objectivity have been developed as devices to improve education, both in school programs with individuals unidentified and in individual instances.

Rice, the founder of comparative educational testing, did not orient his testing toward individual diagnosis and guidance. His concern was primarily with the efficiency of the school as an institution in which children could learn. His point of view was summarized in the opening paragraphs of what was perhaps his most famous paper, "The Futility of the Spelling Grind (Rice, 1897)."

> . . . The first step toward placing elementary education on a scientific basis must necessarily lie in determining what results may reasonably be expected at the end of a given period of instruction. If we have no definite notions in regard to what our teachers ought to accomplish, our ideas must be doubly vague as to how much time need be devoted to each branch. And, as long as this remains unanswered, no well-founded opinion can be given concerning the possibility of broadening the course of study without detriment to the formal branches—the point around which the entire question of educational reform revolves.
>
> Believing that the most rational method of determining what our teachers might be expected to accomplish would lie in discovering what results the more successful ones had been able to obtain, I ventured to undertake a series of researches which I hoped might serve as an initial step toward bringing this problem to a solution [p. 163].

Research methods used by Rice in the final decade of the nineteenth century were not highly sophisticated by present-day standards; yet, as Julian Stanley pointed out in 1966, Rice took into consideration many factors that might be related to success in spelling, including age, intelligence, and foreign language background before coming to the conclusion that the amount of time spent in formal spelling drill was not an important independent variable (Stanley, 1966).

Along with his testing program, which involved thousands of cases and detailed analyses of results, Rice sought to enlist professional educators in the attempt to place education on a scientific basis. In February 1903 he founded the Society of Educational Research, explaining its purposes as follows (Rice, 1903):

> For some years *The Forum* has been instrumental in placing before the public certain facts intended to demonstrate the feasibility of applying the inductive method to the study of practical educational problems. At first, the professional world was sceptical; but it was not long before a number of school people began to take more kindly to the idea; and today there is evidence of a demand of no small proportions for educational facts such as the inductive method alone is able to produce—a demand for exact information, for example, in regard to what results ought to be secured, in the different grades, in the various branches of study, the amount of time that need be allotted to each subject in order that certain specified results may be accomplished, and the methods and conditions that have been proved by experience to be the most efficient in bringing about the desired ends. In other words, it is becoming quite generally appreciated that the results of our various educational experiments should be recorded and systematized in accordance with the dictates

> of science, so that practical school people might be able to formulate their plans of instruction upon a more substantial basis than that of mere personal opinion [p. 118].

Initial membership, especially invited to the organizational meeting at the Murray Hill Hotel in New York, consisted mainly of school superintendents. Rice assumed the directorship of the Bureau of Research which was to carry out research projects aimed at discovering ways of effective instruction. He proposed that the inductive method be applied to determine the grade in which the study of arithmetic should be begun, the role of drill, and the importance of "objective work." In the teaching of English he said (Rice, 1903):

> . . . two questions are preeminent in importance, and embrace the majority of the minor ones. They are:
> 1. To what extent does practice in sentence construction, obtained through the writing of original compositions, the reproduction of stories, exercises in history, geography, etc., etc., contribute to the power of expression and interpretation, and the use of correct English? and
> 2. To what extent does the study of technical grammar contribute to these same ends [p. 124]?

Rice proposed to publish official reports of the Bureau of Research in *The Forum*, of which he was the editor, and to issue supplements as necessary.

In 1904 it was reported about the NEA Convention in Atlanta (Lang, 1904) that

> . . . almost a third of those attending were members of the Society of Educational Research. The progress of this new organization has been truly remarkable. In its one year of existence it has enlisted about 250 members from among the most influential men and women interested in the advancement of the elementary schools. At the annual meeting held in New York University, on February 13, there were represented nine States and the District of Columbia. Letters of regret were received from State, county, and city superintendents, school principals, and college presidents in seven other States. From the side of moral support, the encouragement has been no less gratifying. Many of the members and superintendents not yet directly identified with the Society have undertaken research work in their own schools. Wisconsin is the first State to take official recognition of the movement by tests organized by the State department of public instruction [p. 589].

And a year later (Lang, 1905):

> The membership of the Society of Educational Research is steadily growing. The new idea of educational economics has been discussed in teachers' meetings before large and intensely interested audiences. Educators of large influence have expressed a willingness to lend their support to the movement. In Sweden, England, Germany, and France the working out of different phases of the problem is being taken in hand by professional leaders. The practical teachers everywhere have awakened to the magnitude of the benefits to be derived from the departure for a wiser economy in teaching and the increase of the educational efficiency of the schools [p. 442].

About this time Rice and E. L. Thorndike were in communication with each other, and the January 1905 issue of *The Forum* contains a six page article by Thorndike entitled "The Quantitative Study of Education," in which Thorndike discussed various measurement problems, most of which still concern us:

1. The multivariate nature of intellectual and moral qualities;
2. The lack of commensurate units in measuring the phenomena of education;
3. The errors inherent in educational measurement;
4. The lack of true zero points;
5. The problem of the measurement of change; and
6. Complications in "the measurement of the relationships or dependencies of intellectual and moral facts."

Nevertheless, Thorndike (1905) affirmed his belief in the possibility of educational measurement and its utility in improving educational practices in these words:

> Nor should the peculiar difficulties of quantitative work in education discourage any one from trying to do it. Teachers and administrative officers everywhere will profit from making systematic observations and keeping precise records of school work. They will be less misled by quantitative studies than by guessing. Of course, we shall not advance the knowledge of education by leaps and bounds, for the same reason that would prevent us from so advancing physics or astronomy or chemistry or biology. It requires great natural gifts, special training, and extreme devotion to make great progress in any field of knowledge. No sane person should expect that any change in habits of thought will transform the men and women engaged in educational work into a group of scientific experts. Nor does quantitative work need any such expectation as its justification. All that is asserted is that it is a useful tool for the thinker and is commonly better than the work it replaces [p. 447].

Clearly, Thorndike and Rice saw eye to eye both on the need for tests and on their place in educational research. Thorndike's writings of that period, including his 1904 book, *An Introduction to the Theory of Mental and Social Measurements*, the first statistics book written by a psychologist, heralded the advent of 20th century educational measurement. In the 15 years that followed, Thorndike, his students, and his collaborators were successful in measuring many of the skills and knowledges taught in elementary school: handwriting, reading, language skills, arithmetic, and spelling. School surveys modeled after those of Rice showed the need for improved instructional methods, while the concepts underlying both the standardized test and the test as a diagnostic device for pinpointing individual remedial instruction gradually evolved. It is doubtful whether the great American aspiration of education for all could have been envisaged without the measurement tools forged in their original formats in the workshops of Teachers College and reinforced by instruments developed in dozens of laboratories in various parts of the country.

A few more words about Rice. When he left the editorship of *The Forum* in 1908, the regular section of the magazine called "The Educational Outlook" disappeared, together with all references to the Society of Educational Research. Seven years later, the American Educational Research Association was founded as the National Association of Directors of Educational Research. When the Journal of Educational Research began to appear in 1920 as the official journal of the AERA, E. L. Thorndike and J. M. Rice were listed among the six honorary members, while among the 71 regular members were such measurement stalwarts as L. P. Ayres, B. R. Buckingham, S. A. Courtis, W. S. Gray, M. E. Haggerty, T. L. Kelley, W. A. McCall, W. S. Monroe, H. O. Rugg, M. R. Trabue, G. M. Whipple, and Clifford Woody, many of them veterans of Thorndike's program.

Intelligence testing, inaugurated by Binet in 1905, had an impact considerably different from that of achievement testing. Originally intelligence testing was important only in education and in the treatment of the exceptional child, but later it had an influence on public affairs and business (as well as on scientific psychology). The first Binet scale was devised as an instrument for identifying retarded children, but it was soon found to be equally effective in discovering bright pupils who for some reason or other were not working up to capacity. The Binet method gained widespread acceptance. American translations and adaptations came rapidly. When a committee of psychologists met at Vineland, New Jersey, in May 1917, to devise methods for the psychological examination of recruits for the U. S. Army then engaged in World War I, the chairman of the committee, Robert M. Yerkes, the senior author of one of the revisions of the Binet, suggested that recruit examinations should be individual examinations, consisting of a short series of mental tests which would require approximately 10 minutes for the average subject.

And so it might have been, since all the committee members were well acquainted with the Binet. However, L. M. Terman had brought with him preliminary copies and related statistics of a group test of intelligence developed by his student, Arthur S. Otis. The chief novelty of Otis' method was complete reliance on item types (some of which were new) which could be used in the group situation and which were scorable by an objective key. The new type of measurement was an immediate success and found wide use in the American Army. At last intelligence test records began to be available on hundreds of thousands of cases instead of the dozens and scores of dozens tested with the Binet and its variants.

The original developers of the Army program aspired to apply psychological techniques to help win the war. There was motivation neither to help individuals nor to advance science. However, the Army Alpha was a stunning success and the psychologists were more stunned than anyone. As they doffed their uniforms at the end of the war, they quickly entrained for their home universities, intent upon applying the group test method in education, industry, and vocational guidance. Yerkes and a number of assistants, however, remained in Washington for a number of months, preparing the official report of the examining program and exploring the data for information of interest to scientific psychology. Using punched card methods, they determined the interrelationships of the subtests of the Alpha and the nonverbal Army Beta, global scores on both tests, and the Stanford Binet mental age. They also related test scores, including a combination score based on the Alpha, the Beta, and sometimes a performance scale, to variables such as age, education, military rank, branch of service, civilian occupation, delinquency, and, last but not least, race, and, for the foreign born, country of origin.

It was here that the testers began to believe that their results had important implications for public affairs. In matter-of-fact fashion Yerkes revealed in the official report that of the foreign born, immigrants from Scandinavian and English speaking countries tested high, slightly above the mean of the white draft, while immigrants from Slavic and Latin countries tested low.

Two years later came Carl Brigham's *A Study of American Intelligence*, complete with a foreword by Yerkes. The book involved further analyses of the Army results and included an impassioned plea for restrictive immigration legislation which would keep out the Alpine and Mediterranean hordes which were a threat to American intelligence, while encouraging the immigration of the Nordics who had made this country great.

Those of us who remember Carl Brigham with pleasure tend to be reluctant to revive the memory of a book he later repudiated. Yet, as one of the greatest of the psychometricians of subsequent decades, I think he would have encouraged a discussion of his attempt to influence public opinion by psychological test results.

The book was a part of the zeitgeist of the early 1920s. The theory of Nordic superiority and of the dilution of the superior American stock by the flood of immigration from southern and eastern Europe during the early years of the 20th century was being advanced by a number of popular writers, who relied chiefly on their impressions of what was happening in America. Brigham, like many others in and out of scientific circles, accepted the framework, contributing what he believed to be confirmatory data. He also relied on the notion of strict inheritance of intellectual ability as implicit in the work of Galton and Pearson, and also on the impression of many of the early testers that scores were closely related to inherited ability and even to the goodness of the nervous system.

One of the bits of psychometric folklore which I absorbed many years ago was the fiction that Brigham's book resulted in Section 11 (a) of Public Law 139 of the 68th Congress, the Immigration Act of 1924, which reads as follows:

> The annual quota of any nationality shall be 2 per centum of the number of foreign born individuals of such nationality resident in continental United States as determined by the United States census of 1890, but the minimum quota of any nationality shall be 100.

This was the section under which, for a number of years, immigration from southern and eastern Europe was strictly limited by the United States government, while immigration from northern European countries remained relatively open.

Actually the charts and statistics which Brigham presented in great profusion allegedly demonstrating the decline in American intelligence as the result of immigration from eastern and southern Europe seem to have had only a very minor reinforcing effect on public opinion. This is the opinion of one historian of the period; others ignore him completely. In examining some 459 pages of the Congressional Record reporting the debates on the 1924 Immigration Act, I found no reference at all to Brigham or his results and only one obscure and garbled reference to what was probably Yerkes' official report.

A considerable happier aftermath of the Army testing program was the speedy adoption of the objective item in standardized tests of intelligence, aptitude, and educational achievement, in the measurement of interests and personality characteristics and in classroom tests. Not only did the objective item become an important tool in mass education from the kindergarten through graduate school, but it stimulated the development of psychometric theory, scoring machines, and computer applications, apparatus on which modern psychological and educational testing depends.

I think it can be shown that measurement, by facilitating the fit of the student to his studies, of the job applicant to his work, and of the aspirant to his career has increased human opportunity. In numerous instances testing has revealed unsuspected talent which, when opportunity for development has become available, has resulted in great achievement. We may not have turned many of the mute and inglorious into Miltons, but in some instances at least the mute and inglorious have become physicians, lawyers, engineers, scientists, and educators. Early identification of talent, whether general or specific, has been a major aspiration of the psychometrician, an aspiration not devoid of success.

What about the allegation that testers, by setting artificial standards for admission to colleges and professional schools and for acceptance for government positions and jobs in private industry, have been restrictive, particularly with regard to individuals not in the main cultural stream? Here I think that we, like Carl Brigham, have sometimes been guilty of accepting what seem to be scientific results and of drawing conclusions too rapidly.

When a test has been valid for one segment of a population, we have sometimes assumed that it is generally valid and that a uniform and unvarying cutoff can be applied. It is really only in recent years that we have realized that all test scores, whether they are called aptitude scores or achievement scores or something else, are related to prior experiences, some more, some less. While granting that prior experiences themselves are in many instances qualifying for an educational or vocational experience, we still lack knowledge of how to combine salient facts about the prior history of the individual with test results and other observations so as to permit the best possible decision about the individual, considering his own interest and the interests of the organization and of society.

The answer, which can never be complete but which can be a better answer than hitherto available, would seem to be in continued, dispassionate study of immediate and long range criteria of success in training and in later life. We need more of a multi-multivariate approach to predictive information. In 1975, who will be the best physicians of 1985, and what is the range of behaviors subsumed under the term "best physicians?" What range of talent is needed for an adequate supply of economists, engineers, political scientists?

I am convinced that no complete solution will ever be found. I am sure, however, that the measurement of individual differences has already supplied numerous parts of the solution and that still more parts of the solution are yet to come.

Psychological tests are tools. The inventor of a tool is seldom if ever able to restrict its use to the purpose for which the instrument was first devised. Many tools have a wide variety of uses, while others are highly specialized, but even with a specialized instrument new applications are often discovered.

Whether measurement leads to an increase in educational opportunity depends on administrators of education programs as well as on the psychometricians who devise the tests. In the past 60 years the applications of psychological and educational measurements have increased enormously through the joint efforts of both administrators and test technicians. By and large these applications have improved educational opportunity.

REFERENCES

Brigham, C. C. *A study of American intelligence*. Princeton, N. J.: Princeton University Press, 1923.

Lang, O. H. The Educational Outlook. *The Forum*, 1904, **35**, 577-589.

Increase in Educational Opportunity

Lang, O. H. The Educational Outlook. *The Forum*, 1905, **36**, 429–442.

Rice, J. M. The futility of the spelling grind. *The Forum*, 1897, **23**, 163–172.

Rice, J. M. The Society of Educational Research. *The Forum*, 1903, **35**, 118–129.

Stanley, J. C. Rice as a pioneer educational researcher. *Journal of Educational Measurement*, 1966, **3**, 135–139.

Thorndike, E. L. *An introduction to the theory of mental and social measurements*. New York: Science Press, 1904.

Thorndike, E. L. The quantitative study of education. *The Forum*, 1905, **36**, 443–448.

Today's Testing: What Do We Know How to Do?

A. N. HIERONYMUS
University of Iowa

What do we know how to do—at least reasonably well? This is the basic question of accountability: accountability of the profession of educational and psychological measurement to pupils, parents, other educators, and the general public.

Testing has always been a subject for lively debate and a popular target of critics. While we have profited greatly from dialog with well-informed, perceptive critics, it is doubtful that there is another profession that is as prone to, and capable of self-criticism. In fact, at times the rhetoric might be aptly described as intolerant, destructive, or even masochistic.

What do we know how to do? Everything and nothing. Everything in the sense that it is almost inconceivable that anyone can come up with a problem which has not been or is not being researched. The quantity of the output is overwhelming. A cursory skimming of research publications in measurement and related journals, the tests, the manuals, the position papers, and speeches published in a single year constitutes a formidable assignment; mastery an impossible one. The massive *Research at ETS: Projects and Publications* contains 501 entries; the latest edition of Buros, nearly 1800.

On the other hand, it is doubtful that there are any among us who would claim to have a final, absolute solution to any of the problems that challenge us. Next year always holds the promise of improvement if not of breakthrough, whether the problem is one of conceptualization, implementation, technology, or interpretation.

What criteria can we employ in determining what we do relatively well and relatively poorly? The *APA Standards for Educational and Psychological Tests and Manuals* are relatively adequate as a guide to the design and interpretation of validation research of the more traditional measurement activities. Similar "standards" are often proposed as an inherent part of the presentation of innovative programs. It is possible that additional *external* standards are needed which apply more directly to measurement designed to *compare* the effectiveness of instructional programs, although it is difficult to conceive of any aspect of validation research which Cronbach (1970) did not include in his chapter in *Educational Measurement.* It would seem appropriate to repeat his statement that "validation is the task of the . . . interpreter. Others can do no more than offer him material to incorporate into his thinking [p. 36]."

I have arbitrarily classified the major areas of progress under four headings: Definition of Constructs, Use of Technology, Development of Instruments to Serve a Variety of Purposes, and Use of Measurement in Accountability.

Definition of Constructs

One of the things we know how to do is to look for, operationally define, and experiment with newly conceived variables or old variables renamed or redefined. Most of these would fall into the general classification of interacting independent variables. Precision in defining and measuring situational and interpersonal variables has resulted in slow but steady progress in the explanation and prediction of behavior. Many able researchers are devoting their professional careers to relatively isolated and independent activity with a minimum of institutional or financial support. It is not surprising, therefore, that the list of such variables is long, diffuse, and for the most part unorganized.

The list includes constructs related to achievement motivation, anxiety, aspiration, attention span, creativity, perseverance, parental attitudes, and dimensions of musical aptitude, to name but a few. It includes concepts from Piaget's theories of cognitive development, the variables corresponding to the cells in Guilford's model of intellect, variables in perceptual or cognitive style, and the cognitive, psychomotor, and interest variables in aptitude batteries for specific occupations.

The constructs of today are constantly being refined, interrelated, and incorporated into educational and psychological theory. More and more research is designed to investigate aptitude-situation-treatment-outcome interactions. However, the results of such studies have not been very productive to date, and even the most optimistic appraisal cannot help revealing how much is still to be learned.

Use of Technology

A good many of the things we know how to do have been a direct result of technological developments in optical scanners and computers. These are well-known and need only be summarized briefly.

1. There has been considerable impact on test development. Very complete, complex, and inexpensive analyses of tryout data are almost routinely available to the test constructor. The development of machine-scorable test booklets has opened up possibilities for item types previously regarded as impracticable because of scoring problems (Lindquist, 1968). Technology has also made it possible to process items in which responses are differentially weighted or which require other complex scoring procedures.

2. Entire tests can be assembled and even printed by computer from item banks. The image this procedure conjures up is disturbing to most of us; yet the computer can do many of the routine sorting and screening tasks of a clerical nature. Computer-stored item banks also provide the possibility of better "teacher-made" tests, and a sharing of items for use in various instructional systems.

3. The use of computers makes feasible more efficient standardizations and the preparation of multiple norms. Data on school characteristics stored in the computer memory provide a basis for selecting accurate, efficient samples. Some nationally standardized batteries now make routinely available the choice of such norms as national, regional, Catholic, large city, and local. The possibilities of pupil, school, and item norms is almost unlimited.

Similarly, in programs such as those of the College Entrance Examination Board and the American College Testing Program, equations developed from applications of multivariate analysis provide differential prediction of success which can be used in selection and placement in a wide variety of institutions, subject areas, and circumstances. Particularly intriguing are Novick's (1971) applications of Bayesian methods to prediction. This technology is also appropriate for occupational selection, placement, and training.

4. The use of computers for test processing makes possible the production of an almost unlimited quantity of reports to teachers, counselors, and administrators. In fact, the test user is deluged with information, much of which he may not be in a position to interpret wisely.

What would seem to be crucial in future computer applications are systems of data reduction related to various decision-making processes. Beginnings have already been made in identifying profile discrepancies, but such applications are still in their infancy.

5. Computer technology has also found its way into testing applications in instructional systems. The uses in computer-assisted instruction and in computer-managed instruction are well known. In such systems the computer is used for test processing, record keeping, and in decision making, and may be used directly in the instructional process.

6. The computer has also aided in a wide variety of psychometric applications, both practical and theoretical. The factor analysis of large matrices is routine and economical. Monte Carlo studies have contributed to understanding of sampling distributions. The computer has also made feasible large-scale psychometric research projects such as Project Talent, National Assessment, and the International Project for the Evaluation of Educational Achievement.

Development of Instruments to Serve a Variety of Purposes

Possibly the greatest improvement in measurement practice has come about through the awareness of the need for instruments which are specifically designed to accomplish a single purpose, or at least a limited number of well-defined purposes.

Many nationally standardized tests of achievement are designed, advertised, sold, purchased, and used to serve multiple purposes. It should be added that they have been used for purposes for which they were never intended.

In looking over the current crop of achievement tests, one notices considerable improvement over those in the past, especially in terms of the relevance of content and in the ingenuity of item writers in getting at higher order mental processes. But even a greater improvement would seem to lie in the use of the tests. A large proportion of standardized achievement tests are now administered early in the school year to serve as a guide to the planning of instructional activities. They are thus being used more for formative than for summative evaluation purposes. Tests used this way can reveal strengths and weaknesses or unevenness of development, may lead to further, more specific, diagnostic testing, and hopefully will result in more appropriate educational programming for individual pupils.

Another important emerging development is that of individualized testing. Branching tests, tab tests, and applications of computer-mediated sequential testing are finding their way into practice in a growing number of applications. There are a few instances in which a short screening test is used to assign *S*s to test levels. (In a way this has been done for many years in the administration of the Stanford-Binet.)

Flexibility is also being introduced into the administration of nationally distributed standardized tests. Newer tests reflect the gradual movement of school practices and textbooks away from a strictly graded program in the direction of the implementation of a continuous progress philosophy. It is not unusual to find tests being administered "out-of-level" in EMR classes, in Title I evaluation projects, or in schools whose average level of performance differs markedly from the norm. Such adaptations as have been used to date are far from ideal, but they are more appropriate than the administration of levels without regard to pupil characteristics. The current edition of the *Iowa Tests of Basic Skills* makes provision for the administration of different levels of the tests to different pupils in the same classroom (Hieronymus & Lindquist, 1971). The hoped-for advantage is that each pupil will be provided with a test which is less frustrating, more challenging, more accurate, and more appropriate. There are all sorts of problems in assignment of levels, interpretation, scaling, norming, and communication, but they are not insurmountable. This is just a beginning. The next step is to individualize by subtest and then to build different tests for different subpopulations with different needs. This is something we know how to do. It is time we quit explaining why it cannot be done, and do it.

Perhaps the most spectacular of the changes, at least in terms of activity, is in the development of tests to monitor, modify, and evaluate the effectiveness of certain types of instructional procedures. The procedures tend to be very specific and limited to a narrow range of educational objectives. This implies the use of instruments which a) are built to very carefully prescribed behavioral objectives, b) are sensitive to change *via* intervention, and c) *may* constitute an operational definition of "success" or "mastery."

In general, tests which have been designed with very restricted content specifications to serve a very limited range of highly specific purposes have been referred to as criterion-referenced tests. (Similar tests proposed for use with somewhat different instructional models have been generally known as "mastery" tests.) There are a number of bothersome semantic problems which attend the use of the term "criterion-referenced." The term, "criterion-referenced test," as Glaser & Nitko (1971) point out, "has a somewhat different meaning from the two more prevalent uses of the terms criterion or criterion tests in educational and psychological measurement literature [p. 653]." Indeed, this use of the term is about as far as it can be from the use of the term "criterion-related validation" as is used in the *APA Standards* which implies correlation and prediction. There is also the question of whether the adjective "criterion-referenced" applies to tests, purposes, or interpretations. Popham & Husek (1969) have indicated that a criterion-referenced test cannot be distinguished from the other type by simple inspection; that such tests are distinguished by purpose, construction procedures, specificity of information, generalizability of test performance to domain and use. All of these distinctions are obviously a matter of degree. Furthermore, the construction model for the "other type" starts with the same procedures, the sampling of test tasks from a universe of relevant tasks. Thus, both types could be designated as criterion-referenced in this sense.

Tests of the "other type" to which I have been referring are generally labeled "traditional tests" or "norm-referenced" tests. Contrasting criterion-referenced with norm-referenced tests would seem to imply a critical distinction in terms of interpretation, i.e., a norm-referenced interpretation *vs.* a non-norm-referenced interpretation. The latter is presumably a direct interpretation of some type of raw score, often in relation to a more-or-less subjectively determined raw score cutting point which may be described as passing, acceptable performance, or mastery. Of course, the same types of interpretations may be made of raw scores from the so-called norm-referenced tests, but these interpretations are not often advocated or encouraged. And, as Klein (1970) suggests, norms may be derived from tests labeled as criterion-referenced, to aid in interpretation.

The tests employed in the National Assessment Program, also frequently referred to as criterion-referenced tests, seem to differ substantially from the criterion-referenced tests used primarily for program evaluation. These are not "tests" in the usual sense in that the analyses have been focused on group performance on individual items. Even in national assessment, crude norms are implied, in that comparisons of item *p*s are made between different groups; the performance of a given group is placed in relation to other groups. It would be a small step to deriving norms for states or local systems. For that matter, indices of group performance for individual items have been provided as routine service to schools for a long time for many of the nationally standardized test batteries. Two different kinds of interpretations are generally suggested (Hieronymus & Lindquist, 1971). The first is a comparison with item norms; the second is a comparison of performance with a subjective judgment of how a particular group "should" perform, to which the phrase "criterion-referenced interpretation" might be considered appropriate.

It is realized that the focus on the semantic aspects of terms constitutes nit-picking to a large degree. On the other hand, confusion in terminology and concept appears to be rampant among teachers and administrators, if not among measurement specialists.

It has not been my intention to put down the appropriate use of any of the variations of criterion-referenced tests. Glaser & Nitko (1971), Bloom (1971), Merwin & Womer (1969), and others have presented very thorough and convincing rationales for the need and use of such tests to serve a variety of purposes. The criterion-referenced approach is probably making its greatest contribution in the monitoring and assessment of instructional strategies and outcomes. Applications range from programs such as the IPI project (Lindvall & Bolvin, 1967) and Project PLAN (Flanagan, 1969), which focus on instructional procedures, to National Assessment, which is concerned with description and the relatively long-range monitoring of outcomes.

This approach to evaluation would also seem to have an important place in communication and accountability at the local level. Even in situations in which direct accountability benefits may be questionable, there is the potential for improved communication through the involvement of the lay public and educators in the process of thinking through and operationally defining behavioral objectives.

Use of Measurement in Accountability

If we were to adopt a theme song for 1971 it would likely be entitled, "Accountability—That's Where It's At."

There have been innumerable conferences, speeches, and papers on the subject. Especially recommended are the papers by Lennon (1971), Dyer (1970), Barro (1970), Innes (1971), Stake (1971), the American Institutes for Research volume, *Evaluative Research* (1970), and the papers in the ETS *Proceedings of the Conference on Educational Accountability* (1971). With minor exceptions, my views are congruent with those of Lennon.

The common element in most of these papers is a statement that accountability means different things to different people, and that the basic questions are: Who should be accountable? To whom? For what? And by what procedures is accountability to be determined?

Here are some rather imprecisely stated generalizations which imply some of the issues.

1. Accountability in *some* form is generally acceptable to all parties concerned. As Lennon (1971) said, "Accountability . . . is an idea whose time has come—again, or perhaps an idea whose time is always." School personnel have always regarded accountability as an important responsibility. Most objections are not to accountability *per se* but to some of the instructional and evaluation procedures designed to implement it.

2. The optimum evaluation design for accountability purposes is not often the best design for differentiating instruction and stimulating creative teaching. For years we have been trying to convince teachers that they have nothing to fear from tests, that tests are for their use as aids to better instruction. Emphasis on tests for accountability purposes will almost inevitably contribute to misinterpretation and misuse of tests.

3. An important distinction can be made between group accountability and individual accountability. In its extreme form, group accountability refers to the *joint* accountability of *all* concerned to the *group* of pupils who attend a particular school in attaining a *uniform* set of objectives. For group accountability, the use of existing tests and those under development are probably relatively adequate for the

limited objectives they represent, provided an appropriate design is employed to relate performance to independent variables.

Individual accountability is another story. By this I mean the responsibilities a teacher, counselor, psychologist, and so on, has to an individual pupil on a one-to-one basis to help him to develop his uniqueness. Here the objectives vary and are likely to be long-range. Tests which maximize individual differences both qualitatively and quantitatively are at least partially appropriate, particularly when they are used for formative purposes. This is not to say that the measurement of mastery outcomes is unimportant. Bloom (1971), in particular, has written about the potential contribution which mastery learning may make to self-actualization. But it should be recognized that the contribution of mastery tests to total individual accountability is limited.

4. There is also an important distinction between accountability in terms of product and process. In evaluating product we may employ a fairly straightforward input-output model. But as Wallace (American Institutes for Research, 1970) points out, in evaluating process the basic accountability question is "What happened?", and the focus of attention is on ". . . . what helps, how it helps, and when and where it succeeds or fails [p. v]." This variety of accountability requires assessment by instruments sensitive to short-range change which are used for continuous monitoring. Its basic purpose is to establish cause and effect as directly as possible.

5. In performance contracting a distinction can be made between a "service contract" and a contract for the development of procedures which can be turned over to the school system (the so-called "turnkey" provision). The latter type of contract, which calls for installation and evaluation, should be conceived of and designed in exactly the same way as any other "methods" study, which in fact it is.

There are a great many accomplishments I have not mentioned: better test manuals, better aids to interpretation and use, considerable work in scaling, recognition of social, cultural, and linguistic variability (Holtzman, 1971), among others. If the endeavor to which you are devoting your professional career is among these, please consider it as an accomplishment so well known that its mention is unnecessary.

Summary

In closing, I will try to answer the four questions posed in the program as directly and tersely as possible.

Q: Has the attempt to have testing serve multiple functions with the same test been productive or counter-productive?

A: Productive—if only in the sense that inadequacies in traditional tests have led to the development of tests which will serve the functions for which traditional tests have been ineffective.

Q: Are norm-referenced tests different from criterion-referenced tests?

A: Yes, and speaking for both sides, *vive la différence*.

Q: Should today's standardized tests be used to hold schools accountable for meeting their objectives?

A: The appropriateness of today's tests for the multiplicity of accountability applications is limited. It is hoped that tomorrow's tests will better serve these important functions, but let's not cop out on our obligations to students in order to find something we can objectively account for.

Q: Are there some things we have been doing right?

A: Not in an absolute sense; we have not yet achieved a final mastery criterion.

REFERENCES

American Institutes for Research. *Evaluative research: Strategies and methods.* Pittsburgh: American Institutes for Research, 1970.

Barro, S. M. An approach to developing accountability measures. *Phi Delta Kappan*, 1970, **52**, 196–205.

Bloom, B. S., Hastings, J. T., & Madaus, G. F. *Handbook on formative and summative evaluation of student learning.* New York: McGraw-Hill, 1971.

Cronbach, L. J. Validation of educational measures. In *Proceedings of the 1969 Invitational Conference on Testing Problems.* Princeton, N. J.: Educational Testing Service, 1970. Pp. 35–52.

Dyer, H. S. Toward objective criteria of professional accountability in the schools of New York City. *Phi Delta Kappan*, 1970, **52**, 206–211.

Educational Testing Service. *Proceedings of the Conferences on Educational Accountability*. Princeton, N. J.: Educational Testing Service, 1971.

Flanagan, J. C. Program for learning in accordance with needs. *Psychology in the Schools*, 1969, **6**, 133–136.

Glaser, R., & Nitko, A. J. Measurement in learning and instruction. In R. L. Thorndike (Ed.), *Educational measurement*. Washington, D. C.: American Council on Education, 1971. Pp. 625–670.

Hieronymus, A. N., & Lindquist, E. F. *Teacher's guide for administration, interpretation, and use: Iowa Tests of Basic Skills*. Boston: Houghton Mifflin, 1971.

Holtzman, W. H. The changing world of mental measurement and its social significance. *American Psychologist*, 1971, **26**, 546–553.

Innes, T. C. Measurement, accountability, and humaneness. *Measurement and Evaluation in Guidance*, 1971, **4**, 90–98.

Klein, S. Evaluating tests in terms of the information they provide. *UCLA Evaluation Comment*, 1970, **2**, 1–6.

Lennon, R. T. Accountability and performance contracting. Invited address to the American Educational Research Association, New York City, 1971.

Lindquist, E. F. *The impact of machines on educational measurement*. Bloomington, Indiana: Phi Delta Kappa International, 1968.

Lindvall, C. M., & Bolvin, J. O. Programmed instruction in the schools: an application of programming principles in "Individually Prescribed Instruction." In P. Lange (Ed.), *Programmed instruction*. 66th Yearbook, Part II. Chicago: National Society for the Study of Education, 1967. Pp. 217–254.

Merwin, J. C., & Womer, F. B. Evaluation in assessing the progress of education to provide bases of public understanding and public policy. In R. W. Tyler (Ed.), *Educational evaluation: new roles, new means*. 68th Yearbook, Part II. Chicago: National Society for the Study of Education, 1969. Pp. 305–334.

Novick, M. R., et al. *Applications of Bayesian methods to the prediction of educational performance.* ACT Research Report No. 42. Iowa City: The American College Testing Program, 1971.

Popham, W. J., & Husek, T. R. Implications of criterion-referenced measurement. *Journal of Educational Measurement*, 1969, **6**, 1–9.

Stake, R. E. Testing hazards in performance contracting. *Phi Delta Kappan*, 1971, **53**, 583–589.

Research Methodology for Educational Change

Samuel Messick
Educational Testing Service

In any consideration of research methodology for educational change, two basic questions loom large at the outset—namely, *what* is it that must be changed? and along which dimensions and in which directions should the change take place? A third important question concerning *how* the change is to be produced usually receives most of the attention, but often largely as a practical problem in the engineering of change, an approach that tends to be more concerned with means and procedures than it is with either of the other two issues—that is, either with the nature of the thing to be changed or with the goals or intended outcomes of the change process.

A research methodology provides a systematic way of looking at a problem, a mode of inquiry for eliciting information and evidence leading to understanding and problem solution. Rather than spend the brief time allotted here discussing particular methods or techniques of inquiry and analysis that might be useful in fostering educational change, let us instead focus upon some of the critical problems in the production of change and see what general methodological strategies or "ways of looking" seem called for. First, let us consider the nature of the base from which change is to be mounted—the nature of the educational arena, as it were, especially those features that are typically targeted for change and those features that tend to impede attempts at alteration. Then, we shall turn to a consideration of the basis for change, of the sources of the ideas underlying new plans and new goals for the educational enterprise.

Targets for Change

One of the most critical features of the educational arena is that it constitutes a *system* in the technical sense of a complex set of elements that functions as a whole by virtue of the interdependence of its parts.[1] There are many parties to the system—such as teachers, students, parents, principals, counsellors, communities, superintendents, school boards—and many parts and processes.

In one sense the system is highly dynamic, in that particular individuals involved frequently change or are replaced. Individual students flow through the various grades or levels of the system, for example; teachers are transferred or promoted; and some communities have such a high rate of population exchange that not only is considerable student replacement introduced within grade level but the community's social characteristics are gradually transformed. It seems likely that strategies of intervention and change will have to vary depending upon the rates of such population exchanges—that is, depending upon the degree of interdependence between particular individuals and the organization or system.

In another sense the system is highly stable, in that it displays a huge number of programmatic and behavioral regularities. Most programmatic regularities are intended to produce or to change behavioral regularities. For example, the inexorable programmatic regularity that confronts each student on every school day of every school year with numbers and mathematics is presumably intended to increase in cumulative fashion certain behavioral regularities in the student, as reflected in his repertoire for quantitative thinking and his level of mathematical mastery. The frequency and level or quality of these behavioral regularities, in turn, offer important criteria for judging the extent to which the intended outcomes are being realized.

Not all of the existing behavioral regularities of students, teachers, or other participants in the system are intended outcomes or processes, however. Many of them reflect extraneous consistencies, while others represent unintended outcomes or side effects, which may be either desirable or undesirable. A finding, for example, that teachers ask questions in class at a very high rate (about 50 times as frequently as students do) and that a very large proportion of these questions require straight recall from the students, is a behavioral regularity, to be sure, but is it an intended or a desirable state of affairs?

Since most of the impetus for educational change stems from someone's noticing a regularity that he considers undesirable (or the absence of a regularity he considers desirable), it is not surprising that most efforts at change attack the offending regularity frontally. Failure to appreciate that the regularity is embedded in a system of interdependencies has double-edged consequences. To begin with, the successful modification of a part of the system may have unintended and possibly adverse effects on other parts of the system, thereby not improving matters very much overall. Furthermore, if the network of interdependencies supporting a regularity is overlooked, it cannot very well be countered, so there is little likelihood that circumscribed alterations will be successfully maintained in the face of inertial pressures toward reabsorption of the changed part back into the old pattern. The more things change, the more they stay the same. And the situation is even more intransigent than this, for we are dealing not only with a system but with a *culture*. The existing regularities are determined and maintained by roles and expectations, attitudes and values, tradition and history. Under these circumstances, effective and long-range educational change requires not only individual change and program change but institutional change.

---

[1] Many of the points made in this section are summarized from Seymour B. Sarason's illuminating book, *The Culture of the School and the Problem of Change* (Boston: Allyn and Bacon, 1971). The interested reader is referred to that source for an elaboration of the issues.

Let us briefly consider a case history of a recent intensive effort to modify some perceived undesirable regularities. The case concerns the introduction of the new math into the school curriculum. In reporting the results of a recent observational study, Sarason (1971) summarized the situation as follows:

> . . . for some time before the first Russian Sputnik in 1957 there was a good deal of dissatisfaction with the teaching of mathematics in the public schools. The leadership of this dissatisfaction was primarily in the universities, and the content of the criticism took different forms. But on at least one point there was complete agreement—the way children were being taught math was an unmitigated bore and disaster that very few children could survive either in the sense that they experienced the joy of the world of numbers or pursued mathematics as a career. The Russian Sputnik catalyzed the effort to change the teaching of math, and various new maths were developed in university centers and introduced into the schools. After several years of the new math we observed . . . (its teaching) in a number of classrooms in several school systems. . . . joy is the last word in the English language that one could apply to the children in those classrooms. . . . If our observations and those of others have validity and generality, one would have to predict that the goal of more and better mathematicians and scientists . . . will not be met. If so, we

> will have another sad example of how the more things change the more they remain the same [pp. 19, 46].

As Sarason (1971) points out, this conclusion is perhaps too generous, for not very much has basically changed at all. Why? To begin with, the original diagnosis of the problem focused upon a particular set of undesirable behavioral regularities and their presumed source in the programmatic regularities of the traditional mathematics curriculum; change was to be brought about through substituting a new set of regularities by means of a new curriculum. Very little was said about teachers or the teaching process, nor was the problem formulated in such a way as to require the involvement of teachers as an integral part of the change effort. Furthermore, the impetus for change came primarily from outside the school setting, and the new materials were introduced into the schools, in many cases imposed upon them, with little attention paid to the existing institutional culture or its social and psychological concomitants.

In the settings observed, teachers were trained in the new math in summer workshops in highly traditional fashion as if this were a standard problem of imparting knowledge and developing skills, with little recognition given to the difficulties entailed in having to unlearn highly overlearned ways of thinking at the same time that new modes of conceptualizing were to be acquired. The teachers had unanticipated difficulty in learning the new material to a degree comparable to their grasp of the old curriculum, which itself was hardly at a level that could be characterized as decisive mastery—it was typically more like an uneasy truce achieved after years of struggle. They were obliged to embark on the new enterprise with very little supervision—nowhere near enough to bolster the tentative understanding or dispel the attendant apprehensiveness. Nonetheless, they settled down, as is their wont, to teach in the way in which they themselves had been taught.

One consequence of such an emphasis upon materials and curriculum content is that all those far-reaching goals and intended outcomes concerning the development of positive affect and thinking skills and manpower resources are lost sight of, and the development and delivery of the means becomes a goal in itself. The whole enterprise then tends to be evaluated in terms of the successful accomplishment of the engineering objective of substituting one set of textbooks for another.

There are several implications of all this for research methodology, and one of them relates to the formulation of the problem: the question of how change is to be produced is not only an engineering problem but a research problem. The process of change in each instance should be conceptualized in such a way that it appropriately takes into account the complexities and dimensions of differentiation of the cultural setting in which change is to occur. This means that we need to develop procedures for uncovering the relevant dimensions and relationships, as well as alternative techniques for dealing with them. Since it is unlikely that proposed change will have the same significance for all of the different parties and groupings involved in the setting, the social forces surrounding change are likely to be multiplex. The change process, then, must be differentiated and flexible to be able to capitalize upon existing sources of support, to recruit additional support as needed, or to confront the forces of opposition.

The catch in all this is that we need a considerable amount of information about existing regularities and social forces in order to conceptualize the change process realistically in the first place. The usual reaction to this state of affairs is to urge the external advocate of change to immerse himself in the situation to get a feel for the interacting forces and vested interests before attempting specific interventions. This is all well and good, but it takes time and is rarely done systematically enough to permit generalization to other settings or even to provide a verifiable data base to substantiate what often turns out to be a private view of system functioning. Another tack is to encourage change from within the system by participants who know the territory well. But there is no guarantee that participation *per se* will provide a clear or unbiased view—quite the contrary. Nor can we assume any intrinsic relationships between familiarity with the setting and one's conception of the change process or of the requirements of experimental intervention. Still another approach is to move outside the existing system and attempt to construct a new system under less constrained circumstances. The hope here is that a parallel system offering clear demonstrations of exemplary practice will somehow move the old systems to emulate it, but there are vast differences between creating a new system and changing an existing one.

Techniques for Monitoring Change

What we need is a systematic basis for understanding the system. We need methodologies for objectively determining and describing basic characteristics of system functioning and for evaluating both the changes in those characteristics and the efficacy of the change process in relation to intended and unintended outcomes of the system and of intervention programs.

One possibility might be to introduce into school systems a continuous program of information collection and analysis to uncover existing consistencies and structures and to monitor their perturbations over time. Such an information and assessment program would have to be comprehensive enough to assess a wide range of student outcomes representing intended and unintended consequences of the educational system, as well as characteristics of teachers, students, administrators, programs, and settings that might interact with each other to produce differential results. It would also be particularly important to include some provision for observing and documenting those most critical of all regularities occurring between teacher and student in the classroom. Given the cultural basis of much system functioning, it would be vital to incorporate procedures for assessing the attitudes and values, roles and relationships, perceptions and expectations, and aspirations and goals of the various parties to the system and to attempt to unravel dimensions of consensus and contention permeating the social matrix. With such a multiplicity of components, it is obvious that the program would have to include an analytical methodology capable of multivariate, longitudinal, and interactional comparisons to provide mechanisms for disentangling the threads of multiple covariation. Ultimately such an assessment effort should attempt to go beyond the description of system functioning as it exists to the development of causal models relating prior conditions and processes to outcomes, perhaps through the use of computer simulation techniques, so that potential consequences of alternative action programs could be systematically anticipated.

An intervention project could be introduced into an educational system being monitored in this fashion with prior warning about existing social pressures and with periodic feedback enabling adjustments of the change process. The impact of the intervention could then be systematically evaluated, in straightforward if not routine manner, in terms of changes in measures of behavioral regularities, especially student outcomes.

Such a comprehensive assessment program is an ideal that would require considerable time and effort to realize. It offers the advantages of a wide range of information bearing on administrative and instructional decision making, evaluation, and accountability, and in the long run could serve as a vehicle for continuing research on the functioning of educational systems. As a continuing program, it would also provide an early warning system signaling the appearance of new types of students and new or changing conditions, thereby permitting timely modifications in system functioning to accommodate to new inputs and circumstances. Its major disadvantages are the time and resources required to develop, implement, and maintain it.

There are several alternatives to this strategy, of course, and one of them is of particular interest because of its responsiveness to the time pressures for change. The foregoing research strategy gains much of its power for evaluating and understanding change by introducing interventions in the course of an information collection program. The alternative to be considered next reverses the sequence and introduces information collection in the course of an intervention program.

Time for Change

One of the most serious difficulties in developing acceptable research methodologies for educational change is the widespread adoption of a consistently warped time perspective for viewing change. Over and over again, with exasperating consistency, the time required to initiate real change and to cumulate lasting effects is woefully underestimated. As a result, much educational planning is based upon relatively brief experiences with programs, and decisions are frequently made to initiate programs—or to terminate them—on the basis of short-term evidence, usually obtained from *ad hoc* studies. This can be particularly devastating when a promising developmental program is subjected to summative evaluation prematurely.

Although there are inherent dangers in making long-range policy decisions on the basis of short-term research, in a time of pressing social problems and rapid social change the alternative of basing such policy decisions on long-term research does not appear to be particularly viable either. The problems are simply too pressing and the conditions changing too rapidly for us to postpone action pending the results of long-term research and development efforts that might increase the likelihood of widespread positive effects. Action is needed *now*. But neither can we afford the wasted time and resources of inadequately researched programs that turn out to be ineffective. There are strong pressures not only for action but for accountability.

One strategy that appears to offer some promise in the face of such conflicting pressures is to do both things at once—both initiate the action program and conduct research on the effectiveness and consequences of the action at the same time. Specifically, this approach calls for the undertaking of evaluative research on a continuous basis as an integral part of intervention program efforts. That is, once the action program is provisionally formulated on the basis of the best available knowledge and previous research, we would then proceed to carry out both the program and the evaluative research simultaneously. This would be accomplished by including within the administration of the program provision for collecting information relevant to its evaluation and improvement. The level of the research effort, of course, should be appropriate to the magnitude of the program, to the soundness of the prior basis for the program, and to the potential consequences of program failure. There would thus be no delay in program implementation, but at the same time the research would be as long-range as the program was.

By periodically monitoring program effectiveness, this continuous evaluative research would process data bearing directly on questions of program evaluation and accountability. But it would also provide information relevant for improving the program, as well as for modifying the program to accommodate to new types of students and new or changing conditions. Such continuous evaluative research would thus provide at the program level a rational basis for adaptive action (Messick, 1970; Messick & Barrows, 1972).

Testaments for Change

Let us now turn to questions of the source and direction of change, of the origin of ideas for new plans and new aspirations in education. At first glance it might appear that in a rational enterprise like education these ideas derive from a careful examination of evidence, that they represent direct implications of the findings of research. But let us analyze that notion for a moment. Consider, for example, a body of evidence indicating that a large part of man's behavior is shaped and maintained by its consequences, which are under the control of environmental contingencies.

Consider next a social or educational theorist like Skinner (1971), having a particular conception about the nature of man—"not of a body with a person inside, but of a body which *is* a person in the sense that it displays a complex repertoire of behavior [p. 199]." He might reason from such evidence somewhat as follows: Since man is controlled by his environment and the environment can be manipulated, we should use a technology of behavior to engineer a society having sufficient controlling contingencies to maintain socially desirable and productive behavior. If this means that autonomous man, the inner man, is being abolished, "That is a step forward," according to Skinner (1971, p. 215). "His abolition has long been overdue," he maintains, for "only by dispossessing him can we turn to the real causes of human behavior [pp. 200–201]."

Now consider another theorist viewing the same data but with a different conception of the nature of man—"To say that a man is a person is to say that in the depth of his being he is more a whole than a part and more independent than servile."[2] Such a theorist would reason very differently. He might see behavior control by environmental contingencies as a problem for education, not a solution, and strive to bring the operation of those contingencies into the awareness of students in an attempt to reduce the preemptiveness of conditioned behavior. We might then create, by a process more akin to societal guidance than to social engineering (Etzioni, 1968 & 1970), a society that is not beyond freedom and dignity, one with enough alternative opportunities to challenge the individual to choose, freely and responsibly, those environmental contingencies that will come to maintain his behavior. In this kind of society, autonomous man lives on, for "freedom is found in that kind of interaction which maintains an environment in which human desire and choice count for something (Dewey, 1922, p. 13)."

What is at issue in this example is not the research findings, although there is indeed controversy over the consistency of results in this area with human subjects. What is at issue is ideology. It is not the implications of research results *per se* that are to be implemented in the proposed strategies, it is the implications of research as interpreted or filtered through a particular ideology about the nature of man and society. In this sense, research does not directly determine the aims of educational practice or educational change, nor should we expect it to. It instead serves to refine, to justify, and in its finest moments to challenge directions that are primarily ideologically determined. Its most powerful impact comes on those rare occasions when it stimulates a change in the mediating conceptions, especially when it produces a change in our conception of the human being as a learner.[3]

---

[2] Jacques Maritain, quoted in J. P. Bradley, L. F. Daniels, & T. C. Jones (Eds.) *The International Dictionary of Thoughts*. Chicago: J. C. Ferguson Publishing Co., 1969, p. 465.

Several such conceptual changes have occurred in education in the past fifty years or so, and serious attempts were made to implement the implications of the new conceptions in practice—as when we shifted from the notion of the learner as an empty organism to that of the learner as a dynamic organism, to that of the learner as a social organism, to that of the learner as an inquisitive, stimulus-seeking organism, and so on. One of these conceptual changes occurred fairly recently in early childhood education when research with very young children indicated that they can indeed learn cognitive skills and meaningfully process symbolic and semantic information, thereby shaking our general assumptions about the maturational limitations of young children and their readiness for systematic educational experiences prior to some magical "school age."

The history of these major conceptual shifts is not particularly encouraging, however, for they tend to represent fairly global reformulations in viewpoint while the phenomena they refer to almost certainly require more differentiated treatment. The human learner may indeed function as a social organism, to be sure, but might he not also function as a dynamic organism as well? And as an inquisitive organism? And even as a partly-empty organism? And more as one than another at different times and under different circumstances? Historically, attempts to implement the implications of each new viewpoint have tended to be relatively single-minded, resulting not only in marked pendulum swings in practice but pendulum swings in different directions. There has thus been considerable change on the educational scene over the years but not necessarily very much progress.

---

[3] Jacob Getzels. Paradigms and Practices in the Contribution of Research to Education. Paper delivered at the Project Aristotle Conference, Washington, D. C., December 1967.

Given the central role of ideology in determining aims and action in education, it would seem imperative that we undertake systematic analyses of how ideological dispositions affect action through the medium of judgment and decision making (Dewey, 1964; Stake, 1970). This would require measurement methodologies, perhaps techniques of multivariate analysis or multidimensional scaling, for assessing consistent individual viewpoints about the nature of man and the aims of education and society and for developing an articulated description of each viewpoint in terms of the structure of interrelationships among its component values (Jackson & Messick, 1963; Tucker & Messick, 1963; Tomkins, 1963; Uhl, 1971). An understanding of the structure of different value systems would be particularly important in attempting to trace their interaction in areas of consonance and confrontation. Different ideologies may exhibit substantial degrees of overlap in shared values, for example, while a particular ideology may contain potentially conflicting value components that do not always lead to mutually supporting implications. The personality and environmental correlates of these viewpoints could also be examined for clues as to their origins and the conditions under which they are likely to predominate.

With this kind of strategy, we might then be able to relate differences in educational objectives to characteristics of decision makers, their underlying values, and their situational constraints. By evaluating the empirical consequences of action programs in terms of structures of interrelated values instead of lists of objectives, we would also be better able to appraise in systematic fashion the import of unintended as well as intended outcomes.

To recapitulate briefly, it looks as if the nature of the thing to be changed, the educational arena, is not only a complex system composed of many constituents but a complex culture comprising multiple roles and pluralistic values, and that if we are to understand the functioning of that system—or to change it—we must take into account the interplay of those roles and values in supporting (or subverting) system regularities. We have seen further that the sources and direction of change are also primarily a function of value perspectives and that individual and social values must, therefore, play a central role—often implicitly, to be sure—in educational planning and policy making. The implications of these points for research methodology seem clear. We need methodologies that no longer strive after Max Weber's perverse ideal of a value-free social science, but rather methodologies that are openly value-centered (Gouldner, 1962).

In short, although the methods of empirical social science research are typically employed to guide means-judgments, usually by investigating effectiveness and efficiency, the proposed strategy calls for the application of social science methodology to help understand ends-judgments and to clarify and improve means-ends decisions. If this approach doesn't exactly make valuation a social science, it would at least make the educational applications of social science more valuative (Dewey, 1970; Edel, 1970).

REFERENCES

Dewey, J. *Human nature and conduct*. New York: Holt, 1922.

Dewey, J. Logical conditions of a scientific treatment of morality. In R. D. Archambault (Ed.), *John Dewey on education*. New York: Modern Library, 1964. Pp. 23–60.

Dewey, J. Theory of valuation. In O. Neurath, R. Carnap, & C. Morris (Eds.), *Foundations of the unity of science*. Vol. II. Chicago: University of Chicago Press, 1970. Pp. 380–447.

Edel, A. Science and the structure of ethics. In O. Neurath, R. Carnap, & C. Morris (Eds.), *Foundations and the unity of science*. Vol. II. Chicago: University of Chicago Press, 1970. Pp. 274–377.

Etzioni, A. *The active society*. New York: Free Press, 1968.

Etzioni, A. Essay review of *Crisis in the Classroom*. *Harvard Educational Review*, 1970, **41**, 87–98.

Gouldner, A. W. Anti-minotaur: The myth of a value-free sociology. *Social Problems*, 1962, **9**, 199–213.

Jackson, D. [illegible]sick, S. Individual differences in social perception. *British Journ[illegible]cial and Clinical Psychology*, 1963, **2**, 1–10.

Messick, S. Evaluation of educational programs as research on the educational process. In F. F. Korten, S. W. Cook, & J. I. Lacey (Eds.), *Psy-*

chology and the problems of society. Washington: American Psychological Association, 1970.

Messick, S., & Barrows, T. Strategies for research and evaluation in early childhood education. In I. Gordon (Ed.), *Early Childhood Education, National Society for the Study of Education Yearbook, 1972.* Chicago: National Society for the Study of Education, in press.

Sarason, S. B. *The culture of the school and the problem of change.* Boston: Allyn & Bacon, 1971.

Skinner, B. F. *Beyond freedom and dignity.* New York: Alfred A. Knopf, 1971.

Stake, R. E. Objectives, priorities, and other judgment data. *Review of Educational Research,* 1970, **40**, 181–212.

Tomkins, S. S. Left and right: a basic dimension of ideology and personality. In R. W. White (Ed.), *The study of lives.* New York: Atherton Press, 1963.

Tucker, L., & Messick, S. An individual differences model for multidimensional scaling. *Psychometrika,* 1963, **28**, 333–367.

Uhl, N. P. Encouraging convergence of opinion, through the use of the Delphi Technique, in the process of identifying an institution's goals. Princeton, New Jersey: Educational Testing Service, 1971, Project Report 71–2.

Samuel Messick suggests research techniques appropriate to educational change.

Jack C. Merwin extracts concerns of educational decision makers from Session II.

Discussion

Jack C. Merwin
University of Minnesota

From their titles, one would expect the three papers in this session to be quite distinct and independent. But Bill Coffman knew what he was doing when he put them on this program. He has been vindicated, I think, by the several obvious threads that run through and across the three papers. I shall follow a few of those threads in my comments, including concerns about time and what is tested, and a variety of direct and indirect references to the information needs of educational decision makers.

Concern for time comes out in several ways in these papers. The first and most obvious concern is that the authors have treated testing as it relates to educational change and research in different time frames, and this was by design. Dr. DuBois has given us his historical treatment, which provides a basis for the others to highlight the cyclical nature of concerns in testing. In doing this, he adds another chapter to his 1964 Conference paper dealing with testing from 1115 B.C. to 1905 A.D.

Dr. Hieronymus brings clarity to much of the confusion regarding testing that is found today among educators and the public. Dr. Messick's paper focuses on the future, suggesting how to deal with educational change to make it most effective. He challenges us with a mammoth task, highlighting the complex nature of educational change. In the short time I have had to consider his presentation, I am led to believe it says important things to us, and I am sure that I cannot do justice to it in my brief comments today.

The cyclical nature of our concerns in measurement is another aspect of time that runs through these papers. Two examples illustrate many that are in the papers.

If Dr. DuBois had attributed to a current author the 1905 statement of Dr. Rice regarding "the new idea of educational economics" and "the departure for a wiser economy in teaching and the increase of the educational efficiency of the schools," no one would have questioned it. It supports the statement of Roger Lennon regarding accountability as "an idea whose time has come—again—or perhaps whose time is always."

A second example is Dr. Messick's plea for recognition of the cultural aspects of the multivariate nature of educational change when viewed along with the concerns of E. L. Thorndike, expressed in the latter's 1905 paper, for "the multivariate nature of intellectual and moral qualities" and "the measurement of relationships or dependencies of intellectual and moral facts."

Still another aspect of time that carries through the three papers is the reference by Dr. DuBois to Dr. Rice's interest in the amount of time spent in instruction during the early 1900s, the Bloom mastery approach mentioned by Dr. Hieronymus, and Dr. Messick's reference to premature summative evaluation of new programs.

It strikes me that there may be an analogy here to account in part for Dr. Messick's concern. Dr. DuBois mentioned that in the early days of testing, much test development took place as an avenue to completing a Ph.D. degree. Many recent Ph.D. studies have focused on the effectiveness of various instructional approaches; the fact that it is difficult to enter the job market until the sheepskin is in hand has undoubtedly led to much premature summative evaluation. There is also a high probability that current funding patterns for development activities have led to premature summative evaluations, particularly federal funding patterns.

Let me turn, then, to the concern for what is tested, as treated by the speakers. Dr. Hieronymus, not surprisingly, has provided an excellent and balanced discussion of criterion- and norm-referenced tests. Some people have implied that norm-referenced tests and the many teacher-made tests given in classrooms have been developed with little or no consideration of the content that appears in such tests. As Dr. Hieronymus has noted, this is simply not the case.

With the possible exception of Domain-Referenced Achievement Tests, such as those developed by Hively and his colleagues, it is extremely difficult to find a test with characteristics that have been documented carefully enough to permit defensible generalizations to a large and definable body of behavior.

There is, however, an interesting contrast between instruments for what has been called "success" or "mastery" as operational definitions, and the approach proposed by Dr. Messick. According to Dr.

Hieronymus, the former calls for instruments "built to very carefully prescribed behavioral objectives" that are "sensitive to change." Dr. Messick, on the other hand, calls for a comprehensive assessment program that can be used to monitor both intended outcomes and unintended outcomes, or side effects, which may be either desirable or undesirable. This general idea, advanced by Lee Cronbach some years ago, concerns a number of people when evaluation of highly specific behavior modification procedures is limited to the specific goals of such efforts.

Though posed in a slightly different framework, the monitoring models that appear in the recently published Phi Delta Kappa volume, *Educational Evaluation and Decision Making*, include some of the elements of the type of evaluation Dr. Messick seems to be seeking.

Each of the speakers has touched on the fact that ultimately test scores and other data will be used by someone who must make decisions. Dr. Hieronymus notes that computers can provide large quantities of data, and that teachers, counselors, and administrators can be deluged with information, much of which they may not be in a position to interpret wisely. He also alludes to the student as a decision maker when he mentions the possible value of "mastery" testing for student "self-actualization."

Dr. Messick, in his consideration of ideologies, suggests a strategy that might allow us to "relate differences in educational objectives to characteristics of decision makers, their underlying values, and their situational constraints." There is support for such an approach in the recent findings by Dr. Ralph Straton, now of Sydney University, in a study of techniques for assessing the value of educational objectives as perceived by parents, teachers, and students when the individual child is the focus of the rankings obtained.

Dr. DuBois points out Dr. Rice's early belief that "the motivational method of determining what our teachers might be expected to accomplish would lie in discovering what results the more successful ones have been able to obtain," and possibly there is a methodological clue here. In dealing with decision makers and their need for information, perhaps an intensive study of the selection and use of educational information by our more successful decision makers could help us revamp our information-gathering systems to produce more usable and valuable information. And, this might just possibly make a contribution to the quality of educational decisions across the board.

Discussion

In their final comments, each speaker leaves those who labor in the field of testing and research with some comforting assurance that they will not run out of work.

Dr. Messick poses the challenge for us to work systematically in an extremely complex and many-faceted task that would "at least make the educational applications of social science more valuative." Dr. Hieronymus reminds us that we have not yet achieved a final mastery criterion, and Dr. DuBois tells us that there are many parts missing in the answer to the question of how to generate and combine data to make the best possible decisions, and assures us that no complete solution will ever be found.

At the Luncheon, Samuel B. Gould expresses some cautions to observe in the nontraditional approach to education.

Luncheon Address

The Birth of a Non-Traditional Tradition

SAMUEL B. GOULD
Chairman, Commission on Non-Traditional Study

I.

It is very pleasant to have this opportunity to meet with you at the midpoint in your Annual Conference—or perhaps I should say *our* Annual Conference. The theme, "Educational Change: Implications for Measurement," is an intriguing one. I am in the center of a good deal of educational change myself these days in a very personal sort of way and am still adjusting to it happily. What the implications for measurement are in my own case are too vague to identify, as yet. But I'm sure they exist and I shall find them ultimately.

The president of Educational Testing Service has referred to my comments a year ago when I listed the things I had already begun to miss when I left university administration. All of these were facetious —except one, when I said that I missed "the hitching posts of a lifetime, now suddenly removed, when, in some strange and wonderful way, to be tethered was at the same time to be free to move and act in the world of learning." Bill Turnbull apparently listened very carefully to this and, with his remarkable efficiency and persuasive power, saw to it that I was swiftly tethered again and at the same time left free to move in the educational world. It is a marvelous sort of liberating captivity, paradoxical as such a term may sound, a combination of gentle fetters that never chafe and freedom of thought and action that is never questioned. Only the leadership of ETS could have devised such a non-traditional snare for me, and I am an enthusiastic captive.

It may be some sort of commentary on education's resistance to change that conference luncheon speeches still survive in a world so full of swift and swinging transformations. By now we should be using rock music and improvised, topical lyrics, and flashing lights and psychedelic backgrounds to stir the conferees. But here we are at the same old stand even though we are talking about change. The thought must have crossed *someone's* mind that, as a welcome diversion, a conference without a luncheon speaker lies within the realm of possibility. Such a thought, apparently, is not merely non-traditional, however; it is outright heresy. And heresy in educational circles is still like that in the Dark Ages, never far away from the Inquisition. Still, I can't help thinking wistfully of what a truly non-traditional conference might be like.

You know that one of my current responsibilities is to chair the Commission on Non-Traditional Study. I mention this only because it has a great deal of relevance to the theme of the Conference and to the topic I have chosen for my remarks. By now I assume that you are familiar enough with the origin, the nature, and the work of the Commission to make any extended explanation of it superfluous. At least, I hope so. Let me say only that this group of twenty-six men and women from all over the country is making a completely independent and most careful study of non-traditional study in its concepts, its performance, its promise, and its dangers. Our findings are only now beginning to be made public; they will continue to be made so for more than a year to come; and I trust that they may help to clarify what is now happening in American education.

The Commission has deep interest in and concern about new forms of measurement that non-traditional study will make necessary. These are matters of the closest relevance to your discussions today. It is our hope, therefore, that some of the conclusions of this Conference will give us added perspective and clues for improvement as we examine some of our own major problems.

II.

At first glance, the title, "The Birth of a Non-Traditional Tradition," seems like a purposeful attempt on my part to confuse you. It was not so intended. A moment's thought must have given you the key to what I am about to say. If we are to understand completely why education stands today on such unsure footing with whirlpools of change swirling around it and giving it less and less traditional stability, then perhaps we should consider some of the larger circumstances that create the whirlpools. We should also consider the good and the bad about them, how they are likely to affect education, and whether there are or should be ways of assessing and controlling or guiding these effects. Undisciplined and disorderly change leads inevitably to chaos, even though it may be very tempting as a reaction to earlier rigidities and absurdities. This is true in our total lives as well as in the parts we devote to education. Education, after all, is no more than one of the major components subject to pressures from the larger society today.

We have arrived at a time in the academic world, therefore, when real choices are about to be made. They can be forced upon us or we can direct them. And make no mistake about it, they *will* be forced upon us unless we in education act with more speed, more forthrightness, more initiative, more intelligence, and more ingenuity than we have hitherto displayed. If *we* do not act, others will act for us. The mood of the American people in its assessment of educational need is one demanding a new recognition of and a new adaptation to the forces that shape our destiny as a nation.

What is emerging thus far from this mood is no clear or clarion call. It is a cacophony of protest against far more than current educational philosophy and practice. It is an angry but confused set of confrontations about many matters, a series of clouded realizations that all is not well, a rising tide of mistrust and bewilderment. We see this on every front: the political, the social, the economic, the educational, and even the religious. Revelations of unexplained but at least outwardly dubious dealings in politically high places have shaken the traditional faith and pride of the citizen in his country and her motives. Revelations of the gap between promises and achievement in social progress have embittered whole segments of our population and caused them to feel that our democratic and egalitarian ideals are a mockery. Revelations of some of the methods and goals of modern economic life have reintensified the long struggle of class against class. Revelations of outworn methodologies or obsolescent subject matter or growing disregard of the individual, together with an almost arrogant reluctance even to be questioned, have placed our educational institutions not merely in their traditional ivory tower but beyond the pale. Even the church has come in for its share of accusations, and seems split within itself on matters fundamental to the spiritual life of man.

From all this, in spite of the confusion and bewilderment and sometimes the white-hot anger, a new pattern of life is beginning to show its design. It is not confined to youth, although youth are among its leaders. It is not confined to minorities, although their voices are loud in protest. It is not only the white-collar or the blue-collar worker, the small businessman or the conservationist or the priest. It is suddenly *everyone*—the medieval Everyman of the morality play risen again and seeing the sins that surround him—confronted by fear that what he always thought strong and right and immutable is crumbling before his very eyes; seized with an urge to repudiate what has been allowed to happen; stimulated by an overwhelming desire to find a better way. He sees the glass darkly, but he is not about to wait too long for it to clear before he moves. Nor is he in a mood to worry at the outset as to whether he moves in the right or wrong direction.

Strangely enough, every aspect of today's sense of revulsion against what we have always termed American traditions is founded upon the traditions themselves. We are not witnessing a revolt against American ideals; we are seeing the result of neglect. It is a revolt against the means rather than the ends, the ponderousness and sometimes the pretentiousness of these means, the quality and attitude of the minds who devise and direct these means. This is a point worth pausing over and illustrating.

Egalitarianism is a tradition finding its roots in our constitution and stated again and again as the keystone on which every democratic principle is built and every democratic action is taken. But even after two hundred years this is still an unfulfilled tradition, and its progress toward fulfillment is deliberate in speed and sometimes reluctant in motivation. In education it is represented by the concept of full opportunity, the idea that every man or woman, regardless of age or circumstances, shall have access to as much education as is of benefit. Elementary and secondary education are the most successful examples to which we can point; the land-grant and community college movements represent major successes in higher education. But in the latter there are still great gaps, segments of our population unserved, inequities in the quality of education offered, and a persistent lack of unanimity about how valid the tradition itself actually is. So the current impatience of those who seek more educational options grows more intense. A new insistence is prevalent that additional ways must be found, and swiftly, to bring this tradition close to reality for those hitherto denied.

Protection of our natural resources is another long-standing tradition, not so long as that of full educational opportunity, but long enough. Yet here again the patience of the American people is wearing thin as they see the results of man-made pollution and despoilment all around them. With their own health and those of future generations hanging in the balance, they find it difficult to understand why every corrective move must be so slowly and grudgingly made. Their inevitable conclusion is that non-traditional ways must be found to counteract the deliberateness and force the issue to a speedier solution.

The dignity and freedom of the individual is a highly prized tradition in America. But we live in an age that worships bigness for its own sake and is allowing technology to become not just the maker of tools but the shaper of men. As this occurs, individual man disappears steadily into the mass and becomes more and more a manipulated being, caught in the grip of the mass media, computers, and a huge bureaucratic impersonality. Even in education, where he should least expect it, he finds himself swept into a great collective system that makes him feel puny and of no account; slavishly gathering course hours and credits according to standard patterns that brook few exceptions. And so he shows his rebelliousness in every way he can: in his dress, his speech, his life style, his independently created courses of study outside the institution's curriculum, his contempt for his elders, and his repudiation of everything traditional. Yet all he is seeking is what traditionally has always been deemed desirable and, in fact, necessary: a place in the sun as a person, as an individual.

An informed people, with truth as the touchstone of information, is similarly an American tradition. To be free to speak, to be free to ask questions and get honest answers, to form judgments on the basis of a full background of knowledge—this has always been part of our heritage. But adherence to this tradition is often more honored in the breach rather than in its observance. Truth is partial, at best, and honest persuasion is frequently supplanted by cleverness and even chicanery, whether the task is one of selling a product or explaining a public policy. The confidence of the citizen is shaken when he is repeatedly bilked or cheated or misled. Eventually he becomes suspicious of *all* products or *all* statements. It is small wonder that he is resentful and looks for new ways to guarantee his right to know what is real and true rather than what appears to be.

Out of attitudes such as these—the desire to believe in and hold onto traditional values, the growing realization that these values are still relatively unachieved as goals, the recognition that efforts toward their achievement are slow and forced—out of these has sprung the current vogue of the non-traditional approach to problem-solving and to life itself. The new literature, music, theatre, art, and styles of living all reflect a breaking with the past and a rush to create and live in a different sort of world as quickly as possible.

We are witnessing, therefore, the birth of a tradition to be non-traditional, to move away from the norm, to prize individuality and independence above all else—and perhaps, by so doing, to give new strength and power and meaning and reality to the self-same traditions that have always been part of the American dream. Even in its present nebulous or only partially defined state and even with a great body of hostility surrounding it, the swing toward the non-traditional as a way of life is the largest and potentially most powerful movement in America. It touches everyone in one way or another even when he opposes it, and it influences his thinking and actions whether or not he is consciously aware of the fact. It can be a tremendous force for good or evil, for productive reformation or disastrous revolution, depending on how well and with what good order and speed it can be understood and assimilated in our land.

There is a wonderful sort of exhilaration surrounding such a movement. It is fresh and bold and daring; it represents pioneering and innovation and response to challenge. It has youthfulness about it, the excitement of exploring new vistas for mankind, of being unorthodox, of being fearless in the face of adverse reaction—it is in every sense a crusade. But let me hasten to say that it also has flaws and dangers and pitfalls not to be brushed aside lightly. Some of these may not be the fault of non-traditionalism itself, but they exist nonetheless and should be recognized for what they are.

III.

The first of these flaws or dangers originates in an almost uncontrollable impulse of the American people to seize upon the creation of anything new—an idea, a slogan or catchword, a gadget, a type of theatre or film or literature, a game, even a style of education—and to accept it, adopt it, use it, and imitate it over and over so slavishly as to glut themselves quickly to a point of boredom and eventual distaste. Probably the most obvious example of this can be found in television programming. One successful detective mystery show immediately begets six more lesser imitations; one daytime quiz show spawns a whole school of others; the interest in viewing professional football is turned into a species of maniacal marathon. In theatre and films a few shockers with explicit emphasis on sex and nudity expand into a flood of more of the same. In just about every field, a new idea or new approach is swiftly exploited, blown up out of all proportion to its permanent value, and then is allowed to fade away.

There is frequently a cyclic pattern to all this. Do you remember when "progressive education" had its heyday a few decades ago? Now it is reappearing as the "open classroom," just as highly touted at the beginning, yet already beginning to be questioned for its results as was its original counterpart. One can find examples of other educational innovations, taken up quickly with enthusiasm and much fanfare and then disappearing into oblivion. On the other hand, one can also point to other innovations such as the "work and study" concept, begun quietly almost seventy years ago, carefully developed by a small number of institutions, and only now coming into its own as an educational idea with real merit. My point is that this flaw or danger, whether in non-traditional study or anything else innovative, is not something substantive within itself; it is rather a reflection of how as a nation we react to anything new and appealing.

Another flaw lies in the eagerness of the innovator to be first rather than best. We live in times when it is not enough to perform well. Ways must be found to have everyone know about our performance. The communications media are avid to report exciting efforts of all kinds even when their success is still unproved, and those making the efforts are not at all reluctant about the publicity. What is featured and thus remembered most easily by the public is not necessarily the most important element of what is being tried. All a project needs for attention is novelty, color, promise, and preferably the possibility of controversy. Everyone is in a hurry to get into the spotlight first, even before there has been time for prudent evaluations and judgments. In education this rush is contrary to the very meaning and spirit of the profession, but the temptation is great and many succumb to it.

Still another peculiarity of any non-traditional movement is its tendency, if not watched and checked carefully, to lapse into as firm and intractable a rigidity as the traditions it replaces. This is a subtle phenomenon, so subtle that it sometimes goes unobserved until an ultimate result has been attained directly opposite to the spirit in which the movement was launched. The line between conformity and non-conformity can grow finer as the latter grows stronger and is more the vogue. It can, in fact, disappear altogether when the non-conformity movement is so much in the driver's seat and creates so much social pressure that every non-conformist finds himself conforming because he invariably is constrained to act and think like every other. This can happen in dress, in speech, in food, or in other rather superficial and unimportant areas; eventually it can happen in ideas, which are not unimportant at all. What starts as individual freedom ends up as mass consensus. The culmination of this process, if carried to its logical endpoint, is a species of intellectual ghetto, self-imposed and intentional. It has no walls around it, but it is just as pernicious and confining, since it lives according to the very practices it professes to despise. The fact that it may be well-intentioned in no way lessens the dangers it generates. Unanimity has its place, but when it enters the realm of ideas it is suspect, since it usually has about it the odor of fanaticism.

Mention of fanaticism reminds me of one more danger that lies hidden in a non-traditional movement. This is the aura of revolution with which it sometimes likes to surround itself, often out of sheer impatience with orderly change and occasionally out of a perverse desire to sweep everything and everybody aside and start all over again from the beginning. This is frequently a rather youthful attitude, the attitude of the "now" generation who are contemptuous of history and unconcerned about posterity. They feel there is little, if anything to be learned from the experience of the past. They want quick, even immediate, results in whatever is to be undertaken. They are completely willing to leave to future generations the freedom to straighten out their own difficulties regardless of how these were brought upon them.

We can certainly sympathize with the desire of anyone today, young or old, to see the problems of change met and solved as swiftly as they come upon us. But it is a bit more difficult to agree that all past efforts are valueless, whether in education or anything else, and that whatever we do and have is to be discarded before there is something at least relatively proven in effectiveness to take its place. A non-traditional movement almost by definition is vulnerable to extremism, to untested theories and undocumented practices. When it takes a revolutionary path, it not only endangers quality but repudiates the experience and refuses the assistance of those who can help bring it to a point of stability. Without a sense of stability no system can flourish, even (and perhaps most particularly) a non-traditional system.

As I look back on almost forty years in the educational world, it is clear to me that not only in this world but in the larger one surrounding it we are more and more being guided by a desire to break sharply with tradition. So apparent is this that the atmosphere of revolt against the conventional has become a newly born tradition in and of itself. Education cannot withstand the pressure of this phenomenon any more than any other facet of society can, nor should it try. The fact that non-traditionalism is already on its way to becoming a tradition is simply one more proof of how powerful an influence is the present vastly accelerated pace of change. It is time for education to accept this and to conduct itself accordingly.

IV.

Since my comments thus far have been largely about non-traditionalism in the abstract, let me take my few remaining minutes to talk about non-traditional education more specifically. Obviously all I have been saying about the strengths and dangers of a non-traditional movement applies to education. Non-traditional study represents an inevitable trend. We could not stop it if we chose to, nor should we even make such a choice. But we must see to it that this trend is accompanied by due regard for past achievement that still has relevance, for guarantees of high quality, and for maintaining an atmosphere of stability in the midst of swift change.

What the American people need in education today is not a cutting *back* or a wiping *out* but an adding *to*. We need a broader diversity of subject matter and method if new populations are to be served appropriately. We need more flexible access to education for thousands upon thousands of men and women of all ages previously denied. We need new ways to recognize achievement. But all these needs, and others I could mention if there were time today, are not simply a substitution. They are new dimensions in an already functioning system where much that is good still remains to be saved and adapted to modern urgencies. *Nothing* is to be gained from advocating the complete eradication of the past or present system because of its shortcomings. Even if this were attractive as a theory, it would be folly as a practice. *Everything* is to be gained from coordinating what has been successful with what must now be devised as we enter an era of full educational opportunity. Not only will a better system evolve, but one with surer acceptance by and support from the majority of our citizenry. Such acceptance and support are still a distance away even in the midst of the strong surge toward change.

Non-traditional study still defies truly accurate definition but some of its elements can be readily identified. The most fundamental elements are those of *flexibility* and *individualized learning*. The disappearance of current rigidities in our educational patterns—whether they relate to access, or numbers of courses and credit hours, or types of course modules, or time limits, or whatever else—is now a clear necessity. We are still too far back along the road leading toward rectifying our errors of exclusivity. We are not yet making our process of selectivity one of welcoming students according to individual need, proved by capability and motivation, instead of rejecting them according to arbitrary norms. Our new system, however it may be devised in detail, must be based on individual choice of goals and individualized progress toward those goals.

Insistence upon these two elements alone will serve to increase options, open doors, offer hope to additional thousands of people, and cause sweeping revisions and additions to our traditional practices. It will also point to new necessities in measurement and in counseling not easy to provide but essential to the system's success. The present lockstep can and must be broken and supplanted by a highly personalized approach and a high degree of individual responsibility.

Stemming from these two fundamental elements are others: the new segments of our society to be served; the use of interruption of study to enhance the educational process; the coordination of work and study to provide a total and more realistic learning experience; the non-residential factor as a plus rather than a minus; the recognition and evaluation of experience as part of learning; new areas of subject matter and regrouping of old ones; the use of new communicative devices for instruction; the involvement of agencies other than formal educational institutions and the consequent involvement of whole communities or regions; new financial patterns and responsibilities, together with their implications for government, public and private institutions, and the individual.

I am merely listing these since there is no time today to discuss them. Each one deserves far more research and evaluation than it has hitherto received. All these elements have a certain amount of historical background already, since none is totally untried. But factual data about them are incomplete, and conclusions are still clouded with doubt. We need to know more, and quickly, if these elements, individually and severally, are to make a full contribution to the strengthening and to the acceptability of non-traditional study.

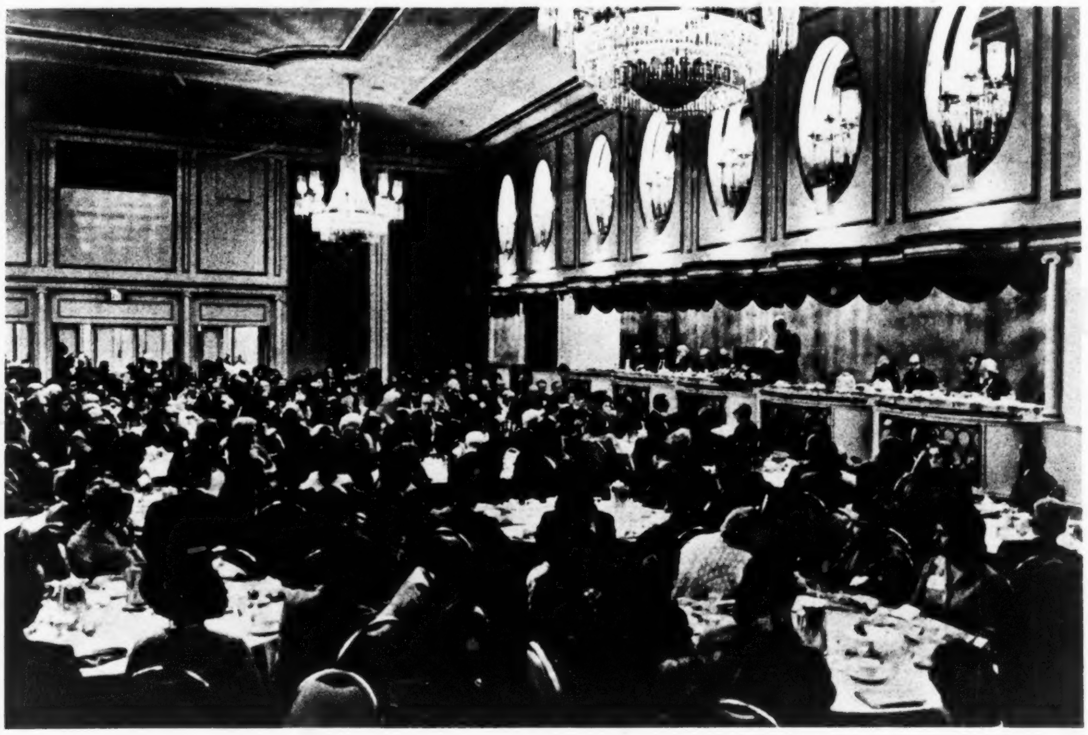
A great weight of responsibility is bound to fall upon the measurement sector, with which this conference is so deeply concerned. Our topics for discussion today show our full awareness of the problem that pressures toward educational change have put before us. But we must have more than awareness. We must have a sympathetic and well-documented understanding of the trend toward the non-traditional. We must have, in addition, a willingness to adapt to that trend with a boldness that involves us as partners in the new thinking and the new action. As we move into uncharted territory, we must do so with optimism—not optimism in the abstract, but the sort that engenders the most ingenious, the most imaginative, the most creative talents we can muster.

Everything else in non-traditional study will rise or fall according to our ability to measure for new and different capabilities of the individual, to guide him wisely, and then to evaluate his achievements by appropriate standards. And since non-traditional study will inevitably grow with full recognition of this measurement need, accountability for the rise or fall in its total fortunes will be placed at our door. We can regard this as a threat or as the most challenging opportunity we have ever discovered. We must decide now which it is to be. I have not the slightest doubt as to what that decision will be and, I suspect, neither do you.

William W. Turnbull (right) gives Lee J. Cronbach the 1971 ETS Award for Distinguished Service to Measurement.

Anne Anastasi, Fordham University, chosen as Chairman of the 1972 Invitational Conference, receives congratulations from ETS President Turnbull.

Guests and Speakers



Session III

New Directions in Education: The Challenge to Measurement

Jerome Taylor urges early intervention in childhood education.

Exploring Selective Problems and Approaches to Early Childhood Education

JEROME TAYLOR
University of Pittsburgh

In this paper, I have opened the sluices of my imagination and experience to consider the case for early educational intervention. I shall argue that most intervention efforts are initiated too late and are focused too narrowly and that most existing approaches suffer from insufficient interfacing with other disciplines.

The Case for Early Intervention

The case for preschool intervention has been amply made and has motivated the proliferation of several curricular approaches (Klein, 1971). The differential effectiveness of these models notwithstanding, some have argued that intervention during the formally designated preschool years of three to five is still too late. Perhaps we can examine some of the reasons why.

There is general consensus that IQ performance differences among occupational and socioeconomic groups begin to emerge between the second and third years. Bayley (1965), measuring IQ performance through 15 months, and Hindley (1965), measuring IQ at 18 months, found no significant differences a function of socioeconomic membership. Schaefer (1969) found that disadvantaged black males at 14 months did not have low Bayley scores, and Francis-Williams and Yule (1967) did not find a meaningful relationship between socioeconomic status and Bayley Infant Mental Test scores between one and 15 months of age.

In contrast to the findings just cited, Hindley (1965) found clear-cut instances of differences among socioeconomic groups at three years. So did Van Alstyne (1929). Moreover, normative data for the Stanford-Binet indicate that IQ scores of children from different occupational groups were as large for children between two and five and one-half years as they were for later age groups (Terman, 1937, cited by Schaefer, 1970). Together, these studies intimate the emergence of group performance differences between two to three years of age.

Given this brief and only suggestive review, the absence of group IQ differences during the first two or three years would seem to justify our efforts toward early intervention. The question now is, how early is early? To my mind, intervention programs that start at birth, one, two, or three years of age are not early enough. Let us examine the reasons why.

The Period of Pregnancy

Several epidemiologic studies (Kawi, 1959; Kawi & Pasamanick, 1959; Pasamanick & Knobloch, 1958) have indicated the relationship between prenatal and perinatal complications and later reading disorders. Other studies (Pasamanick, 1956; Rogers, Lilienfield & Pasamanick, 1955; Stott, 1962, 1965) have established a relationship between complications of pregnancy and delivery and later behavior disorders. Pasamanick (1956) found, for example, that no less than one-fourth of his white sample and one-half of his black sample of behaviorally disturbed children had been exposed to one or more obstetrical complications. Wolff (1967) reported patterned relationships between perinatal factors and later behavioral symptoms of overactivity, restlessness, and distractability. The relevance of these studies for early and later intervention efforts cannot be denied. Let us turn now to a brief review of those factors which seem to mediate prenatal and perinatal complications.

On the one hand, a wide range of infectious, metabolic, and dietary factors are implicated. The list would include maternal rubella (Brown, 1967), mumps (Rhodes, 1961), measles (Hagströmer, 1948), smallpox (Moloshok, 1965), chickenpox (Abler, 1964), and poliomyelitis (Swarts & Kercher, 1954). All these infectious diseases may produce congenitally deficient or malformed infants. Metabolic disorders such as hypothyroidism (Man, Shaver, & Cook, 1958) and diabetes (Horii, Watanbee, & Ingalls, 1966), and other problems such as drug addiction (Claman & Strang, 1962), smoking (Simpson, 1957) and prescribed drugs such as tranquilizers, diet pills, antibiotics, and sedatives (cf. Ferreira, 1969) may all have untoward effects upon the developing fetus.

On the other hand, psychogenic stress may adversely affect maternal physiology and have harmful consequences for the fetus. Davids, DeVault, & Talmadge (1961) found that women who later experienced complicated deliveries or gave birth to abnormal children had significantly higher anxiety scores than women with normal deliveries. Essentially the same finding was reported by Davis & DeVault (1962), McDonald & Christakos (1963), and McDonald & Parham (1969). Zemlick & Watson (1953) observed that the more mothers rejected the fetus, the stormier the pregnancy was and the greater the delivery adjustment. Blau, Slaff, Easton, Welkowitz, & Cohen (1963) found a positive correlation between maternal rejection during pregnancy and the incidence of prematurity. Examining the relationship between maternal attitude, as measured by the Parent Attitude Research Inventory, and prenatal or perinatal complication, as determined by obstetrical examination, Prechtl (1963) found that seven of eight mothers delivering neurologically abnormal babies were either rejecting or overanxious in attitude. While the exact linkages connecting psychogenic stress, physiological reactivity, and perinatal complications are not entirely clear, the findings seem suggestive enough to warrant our concern.

Since prenatal and perinatal difficulties complicate the job to be done later by educators, should they not be more involved in interfacing endeavors with public, medical, and mental health personnel?

From Birth to Three

At this time, I would like to consider the conjunctive importance of physiological, social-emotional, and cognitive variables during the first three years. Not uncommonly, early intervention approaches stress one or perhaps two of these variables but seldom integrate all three. I will argue for the importance of all three variables and their systematic inclusion in early intervention efforts. Naturally, my treatment of this position can be suggestive only and not exhaustive, because of the time limitation.

PHYSIOLOGICAL VARIABLES

A number of investigations have been concerned with early patterns of physiological and behavioral reactivity to multisensory and stress stimuli. Birns (1963, 1965) found that babies could be differentiated within the first five days of life in terms of their reactivity to external stimuli. Babies who, for example, gave a low intensity response to one stimulus tended to give a low intensity response to all stimuli. Included in the stimulus group was a loud tone, a soft tone, a pacifier, and a cold disc applied to the infant's thigh. While the necessary data are not yet available, one wonders about the implications of these findings for differential approaches to infants manifesting extremes of behavioral reactivity. Will the child who is under-responsive require a more active and intrusive approach to cognitive stimulation? Will the same approach have injurious consequences for the child who is over-responsive? We can only answer these questions when they are raised. These questions will be raised only when we begin to think more systematically about potential linkages between physiological and cognitive variables.

Approaching the same subject of early differences in physiological orientation from a syndrome rather than a reactivity perspective is the work of Prechtl (1961, 1963). He identifies several distinct syndromes of minimal brain damage, two varieties of which are discussed. He speaks of the hypokinetic infant who often is hypotonic, drowsy and apathetic, sucks poorly and slowly, and reacts weakly to various stimuli. In contrast, he discusses the hyperkinetic infant who often is hypertonic, startles easily, changes suddenly and unpredictably from one state to another, and exhibits a low frequency but high amplitude tremor. Should the intervention approach be the same for both types of infants? I am reminded here of Gordon Allport's aphorism that it is the same fire that melts the butter that hardens the egg, which I translate as, it might be the same intervention that has distinctively different consequences. I feel we need to look much closer to see if this is so.

COGNITIVE VARIABLES

I am of the opinion that most early intervention programs conceptualize the domain of cognitive growth far too narrowly. Kagan (1971) to my mind has made some evocative recommendations on how to conceptualize this vital area of development. He recommends differentiating the domain of cognitive development into cognitive units and cognitive processes. Let us now examine briefly the component aspects of these units and processes.

Kagan (1971) discusses five types of cognitive units. First, he speaks of *schema*, the central representation of event components which, in early infancy, are sensory representations of the infant's actions toward objects. Second, he calls attention to the development of *images*, the iconic mode, the basic elaboration or holistic representation of existing schema. Third, the emergence of *symbols* refers to the association of arbitrary representations, such as the label "animal," to specific events, like the picture of a bear. Fourth, Kagan speaks of the development of *concepts* where, for example, the child is able to discern the criterial meanings of "animalness." Finally, cognitive units include *rules* which stipulate the relationship between concepts (for example, '*rocks* are *hard*') or transpose two or more concepts into a new one (for example, multiplying two numbers to produce a third).

Cognitive units are conjoined with cognitive processes to solve problems. Kagan (1971) enumerates five different types of cognitive processes which are summarized here. To solve a problem, one must first *comprehend* the nature of the problem. This involves the engagement of selective attention and informational decoding. To see the problem through to solution, its salient and critical aspects must be stored in *memory*. Following the tasks of comprehending and remembering, the child must then *generate hypotheses* by inventing alternating ways of solving the problem. Next is the job of *evaluating* the appropriateness and accuracy of competing solutions to the problem. Finally, the child applies the evaluated idea or rule to the specific problem and through *deduction* arrives at its solution.

Few early intervention programs, to my knowledge, deal in a balanced and comprehensive way with fostering growth developments of cognitive units *and* cognitive process. There might be, for instance, a real concern for developing the concepts of "inside," "outside," and "around," but there is often missing the effort to deploy these concepts in the service of problem solving. Vocabulary building, as another approach, is essential, but should not remain isolated from problem solving activities. Few early intervention efforts are keyed systematically to address and strengthen component aspects of comprehension or memory, hypothesis generation, evaluation, and deduction. For early intervention programs to make a difference, it seems to me that we *must* extend our conceptual analysis of cognitive development in the direction discussed.

We must eliminate the normative utilization of IQ tests on the grounds that they are insensitive measures of cognitive development. The complex dimensionality of cognitive units and cognitive processes is only globally and inadequately sampled, if at all, by existing infancy scales. Only by attempting a more fine-grained measurement of cognitive components can we hope to particularize our approach to specific strength and problem areas of the child. Such is the present challenge to measurement technology. I might also add that without more finely honed measuring instruments for cognitive, physiological, and social-emotional development, we will soon find ourselves defending early intervention from a very solifidian position: Faith alone and not data justifies it all.

SOCIAL AND EMOTIONAL VARIABLES

Powerfully influencing the development of cognitive organization during the first three years are important emotional and social processes. Murphy and her collaborators (1962) have described the emergence of highly individualized patterns of coping with novelty and defending against stress during this period. Important also is the child's growth toward autonomy (Murphy and others, 1962; Sander, 1962, 1964), self-confidence (Murphy and others, 1962), curiosity and exploratory orientation (Ainsworth, 1963), and internal locus of control (cf. Stevens, 1971). The conflation of these growth patterns affects the quality and quantity of the child's engagement with his cognitive environment and, as such, must receive considered attention in any early intervention program.

We are now beginning to grasp how some of the above developments can be facilitated. For example, contingency learning within a general operant framework seems to be a component in developing internal locus of control; secure object attachment seems to be an important way station toward the development of autonomy, self-confidence, curiosity, and exploratory orientation; maternal responsiveness seems important in mediating the infant's adaptive response to stress.

What we have understood and dealt with less adequately is the paracultural* context within which maternal behaviors take place. This context has a past and a present. In terms of the latter, the

---

*I use the term paracultural to avoid some of the negative connotations of terms like "subcultural." The prefix "para" is used to mean *coordinate with* or *developing alongside*.

mother's social system includes the husband or a paramour, other children and the extended family, the immediate neighborhood and the broader community. Any subset of these relationships may importantly influence maternal attitudes. The presence of other children within a context of few resources, a sense of powerlessness in the face of overwhelming need, and the restriction of alternate solutions to persisting problems are all reality impingements which shape maternal orientations. We should never let these realities slip from our view, our concern, and our reach.

I mentioned that maternal behaviors have a past as well as present context. Let's take an example from the black paraculture. There is a feeling among many black mothers and grandmothers that children should not be spoiled. This translates often as admonishment against holding a baby too long or as instructions not to respond immediately to its needs. While these attitudes are multiply determined, I feel they do have a historical past. In the days of the subhastation, the public slave auction, black mothers never knew if they would be with their children from one day or one month to the next. One adaptive consequence of this uncertainty might have been the discouragement of focalized attachments. The prior adaptiveness of this behavior must be re-examined now within the context of present realities. There are no doubt a number of child rearing orientations that articulate with the past. To reorient them requires a knowledge of and sensitivity to that past. Such is an important challenge to our future intervention efforts.

REFERENCES

Abler, C. Neonatal varicella. *American Journal of Diseases in Children*, 1964, **107**, 492-494.

Ainsworth, M. D. The development of infant-mother interaction among the Ganda. In B. M. Foss (Ed.), *Determinants of infant behavior II*. New York: Wiley, 1963. Pp. 67-104.

Bayley, N. Comparisons of mental and motor test scores for ages 1-15 months by sex, birth order, race, geographical location, and education of parents. *Child Development*, 1965, **36**, 379–411.

Birns, B. Constancy of response to internal stimuli in the neonate. Unpublished doctoral dissertation, 1963, Columbia University.

Birns, B. Individual differences in human neonate's responses to stimulation. *Child Development*, 1965, **36**, 249–256.

Blau, A., Slaff, B., Easton, D., Welkowitz, J., & Cohen, J. The psychogenic etiology of premature births: A preliminary report. *Psychosomatic Medicine*, 1963, **25**, 201–211.

Brown, R. S. The expanded rubella syndrome. *Current Medical Digest*, 1967, **34**, 1409–1414.

Claman, A. D. & Strang, R. I. Obstetric and gynecologic aspects of heroin addiction. *American Journal of Obstetrics and Gynecology*, 1962, **83**, 252–257.

Davids, A. & DeVault, S. Maternal anxiety during pregnancy and childbirth abnormalities. *Psychosomatic Medicine*, 1962, **24**, 464–470.

Davids, A., DeVault, S., & Talmadge, M. Anxiety, pregnancy, and childbirth abnormalities. *Journal of Consulting Psychology*, 1961, **25**, 74–77.

Ferreira, A. J. *Prenatal environment*. Springfield: Charles C. Thomas, 1969.

Francis-Williams, J. & Yule, W. The Bayley infant scales of mental and motor development: An exploratory study with an English sample. *Developmental Medicine and Child Neurology*, 1967, **9**, 391–401.

Hagströmer, A. Two cases of congenital malformations after exposure to measles in early infancy. *Acta Pediatrics*, 1948, **35**, 242–246.

Hindley, C. B. Stability and change in abilities up to five years: group trends. *Journal of Child Psychology and Psychiatry*, 1965, **6**, 85–99.

Horii, K., Watanabe, G., & Ingalls, T. H. Experimental diabetes in pregnant mice. Prevention of congenital malformations in offspring by insulin. *Diabetes*, 1966, **15**, 194–204.

Kagan, J. Cognitive development and programs for day care. In Edith H. Grotberg (Ed.), *Day care: resources for decisions*. Office of Economic Opportunity, U. S. Government Printing Office, 1971. Pp. 135–152.

Kawi, A. A. & Pasamanick, B. Prenatal and paranatal factors in the development of childhood reading disorders. *Monographs of the Society for Re-*

search in Child Development, 1959, **24**, No. 4.

Kawi, A. A. Prenatal and paranatal factors in the development of childhood reading disorders. *Monographs of the Society for Research in Child Development*, 1959, **24**, No. 3.

Klein, J. W. Planned variation in head start programs. *Children*, 1971, **18**, 8–12.

Man, E. B., Shaver, B. A., & Cooke, R. R. Studies of children born to women with thyroid disease. *American Journal of Obstetrics and Gynecology*, 1958, **75**, 728–741.

McDonald, R. L. & Parham, K. L. Relation of emotional changes during pregnancy to obstetrical complications in unmarried primigravidas. *American Journal of Obstetrics and Gynecology*, 1964, **90**, 195–201.

McDonald, R. L. & Christakos, A. C. Relationship of emotional adjustment during pregnancy to obstetric complications. *American Journal of Obstetrics and Gynecology*, 1963, **86**, 341–348.

Moloshok, R. E. Viral and parasitic diseases. In J. J. Rovinsky and A. F. Gottmacher (Eds.), *Medical, surgical, and gynecological complications of pregnancy, 2nd ed.* Baltimore: Williams and Wilkins, 1965.

Murphy, L. B., & Collaborators. *The widening world of childhood.* New York: Basic Books, 1962.

Pasamanick, B. The epidemiology of behavior disorders in childhood. *Research Publication of the Association of Nervous and Mental Disease.* Baltimore: Williams and Wilkins, 1956.

Pasamanick, B. & Knobloch, H. The contribution of some organic factors to school retardation in Negro children. *Journal of Negro Education*, 1958, **27**, 4–9.

Prechtl, H. F. R. Neurological sequelae of prenatal and paranatal complications. In B. M. Foss (Ed.), *Determinants of infant behavior I.* London: Methuen & Co., Ltd., 1961. Pp. 45–46.

Prechtl, H. F. R. The mother-child interaction in babies with minimal brain damage. In B. M Foss (Ed.), *Determinants of infant behavior II.* 1963. Pp. 53–59.

Rhodes, A. J. Virus infections and congenital malformations. In the International Medical Congress, Ltd. (Ed.), *First International Congress on congenital malformations*. Philadelphia: Lippincott, 1961. Pp. 106–116.

Rogers, M. E., Lilienfield, A. M., & Pasamanick, B. Prenatal and paranatal factors in the development of childhood behavior disorders. Supplement to 102 *Acta Psychiatry and Neurology*. Oslo: E. Munksgaard, 1955.

Sander, L. W. Issues in early mother-child interaction. *Journal of the American Academy of Child Psychiatry*, 1962, **1**, 141–166.

Sander, L. W. Adaptive relationships in early mother-child interaction. *Journal of the American Academy of Child Psychiatry*, 1964, **3**, 231–264.

Schaefer, E. S. Home tutoring, maternal behavior and infant intellectual development. Paper presented at the meeting of the American Psychological Association, Washington, D. C., September 1969.

Schaefer, E. S. Need for early and continuing education. In V. Dennenberg (Ed.), *Education of the infant and young child*. New York: Academic Press, Inc., 1970. Pp. 61–82.

Simpson, W. J. A preliminary report on cigarette smoking and the incidence of prematurity. *American Journal of Obstetrics and Gynecology*, 1957, **73**, 808–815.

Stevens, M. Cognitive and cultural determinants of early IE development. Paper presented at the meeting of the American Psychological Association, Washington, D. C., September 1971.

Stott, D. H. Abnormal mothering as a cause of mental subnormality-I. A critique of some classic studies of maternal deprivation in light of possible congenital factors. *Journal of Child Psychology and Psychiatry*, 1962, **3**, 79–91.

Stott, D. H. Congenital indications in delinquency. *Proceedings of the Royal Society of Medicine*, 1965, **58**, 703–706.

Swarts, C. L. & Kercher, E. A fatal case of poliomyelitis in a newborn infant delivered by cesarean section following maternal death due to poliomyelitis. *Pediatrics*, 1954, **14**, 235–237.

Van Alstyne, D. The environment of three-year-old children: factors related to intelligence and vocabulary tests. *Teachers College Contributions to Education*, 1929, No. 366.

Wolff, S. The contribution of obstetrical complications to the etiology of behavior disorders in childhood. *Journal of Child Psychiatry and Psychology*, 1967, **8**, 57–66.

Zemlick, M. F. & Watson, R. I. Maternal attitudes of acceptance and rejection during and after pregnancy. *American Journal of Orthopsychiatry*, 1953, **23**, 570–584.

Kenneth B. Hoyt states the case for expanding vocational-technical education.

Broadening Education Through Vocational and Technical Training

KENNETH B. HOYT
University of Maryland

Introduction

The Fifth Report of the National Advisory Council on Vocational and Technical Education (1971) says, in part:

> There is an educational consumer revolt developing in our land today . . . Public officials responsible for education, both elected and appointed, need to be reminded of Alexander Hamilton's statement, "Here, Sir, the people govern! [p. 3]"

In making this statement, the Council has reference, of course, to the fact that a record number of school bond issues have failed in the last two years. They also have reference to the growing need for what USOE Commissioner Sidney Marland has labeled "career education."

It would not seem difficult to verify the fact that something is wrong with our public schools. The high school dropout rate still exceeds 25 percent (U. S. Bureau of the Census, 1970). Nationally, close to 80 percent of all secondary school students are enrolled in either a college preparatory or a general curriculum designed to ready them for college attendance. Yet, we know that it is unlikely that more than 17 percent of these students will ever receive a college degree. We know further that to produce many more baccalaureates than this would create a still larger oversupply of college graduates in the labor market (U. S. Department of Labor, 1971).

At the community college level, there was evidence in 1967 to indicate that, while three in four students are enrolled in a college transfer program, only one in four would ever receive a baccalaureate degree (Collins, 1967).

When in both high schools and community colleges we find three in four students preparing to do what more than three in four students cannot do, it seems evident that something is wrong.

Record numbers of persons are enrolled in the four-year colleges and universities of our nation. Hundreds of thousands of such students do not really know why they are there. Someone told them to "go to college and you will find yourself." The degree of student unrest, disenchantment, and discontent on our college campuses indicates that either this quest has not been successful or that students are dissatisfied with the conclusions they have reached. Thousands of such students are learning a very great deal about how to lead "the good life" but almost nothing about what they could do to pay for it. The days of an automatic guarantee of a job for all who graduate from college are past. Today's college students know this full well.

Thousands of words have been written and spoken about the need for more persons to be enrolled in public school vocational education programs at the secondary and at the postsecondary school level. Yet, with current enrollment trends, it has been estimated that vocational education will supply only 24 percent of the projected labor market needs in 1975 as opposed to supplying 31 percent in 1971 (American Vocational Association, 1971). The integration and acceptance of occupational education as part of American education has not taken place in the minds and thoughts of those who control our educational system. Those who enroll in vocational education are still regarded by most persons as second-class citizens who have made second-best choices.

As a country founded under the philosophy of a work oriented society, our nation is in deep difficulty. Federal legislation, viewed as a solution to these difficulties during the decade of the 1960s, has proven insufficient. The promises of the 1960s demand performance in the 1970s. The fifty million American families for whom work is necessary, but college inappropriate, demand alternatives that will offer them status, security, satisfaction, and success. They cannot and will not wait much longer.

The roots of the problem

Four roots of the problem, each of which is intertwined with the remaining three, exist. All four roots must be extracted and destroyed.

One of these roots consists of the false societal attitude that worships a college degree as the best and surest route to occupational success. This attitude is as dangerous as it is false. When less than 17 percent of the population can attain what close to 100 percent of the population have been led to believe is desirable, it is inevitable that the majority must be dissatisfied with their lot. A viable democracy cannot afford to view 83 percent of its population as "second-class" citizens. Unrest is sure to exist when only one-third can try and less than one-sixth attain what more than nine-tenths regard as the optimum pattern of education as preparation for work.

A second root is found in the generally erroneous application of a basically sound educational principle; namely, that more education can lead one to be better prepared for work. The principle is sound only if one recognizes the equal validity of the principle that any worker can be overtrained for the work he does. Stated another way, the optimal amount of education required as preparation for work will vary widely from occupation to occupation, and from person to person within an occupation. Failure to recognize that expertise can be attained and should be rewarded in any occupation is an essential accompaniment of this erroneous application.

A third root is present in the gradual but persistent erosion of the work ethic as a basic value in American society. Given equal natural resources and ingenuity of its people, the expected productivity per unit of population associated with any society in Western civilization bears a direct relationship to that nation's commitment to the work ethic. No great civilization in the history of mankind has survived long after it abandoned this basic commitment. There is little reason to believe that the United States of America is the exception to this historical pattern. Yet we know the classical Protestant work ethic does not exist among many citizens today. It has been eroded by some of the very factors that make more work possible. The rapidity of occupational change with the resulting instability of occupations, the increased complexity of occupations with the resulting "slicing" of jobs into smaller and smaller components, the movement toward a "throw-away economy" producing products that are never intended to be repaired, with the resulting loss of a sense of craftsmanship, and the increase in geographic fluidity of the sources of employment with the resulting instability in family structure, have each contributed to this erosion.

The fourth root lies in the failure of our American educational system to place proper emphasis on education as preparation for *making* a living as opposed to education as preparation for living. In practice, the prime operational purpose of the classroom teacher seems simply to be readying students for more education. It is time to halt, as Dr. Rupert Evans has said, a "school for schooling's sake" emphasis in American education. Education as preparation for *making* a living must be recognized as an important part of preparation for living.

Barriers to progress in education

American education is now being charged with destroying the four roots identified in the preceding section. The prime substantive solution, to be presented in the following section, lies in installing a career education emphasis at all levels of education. Unless certain basic educational barriers are removed, the career education emphasis of today can be expected to have no more lasting or effective impact than the life adjustment emphasis prevalent 20 years ago. Space will not permit an adequate discussion of these barriers, but it is essential that each be identified.

The first barrier consists of the false worship of college degrees and college credits in the organizational, administrative, and fiscal policies of American education itself. The absolute necessity for college degrees and college credits should be eliminated as a requirement for progress as an educator, as well as in many other fields. Alternative routes to the acquisition of competence must be recognized, utilized, and rewarded.

The second barrier consists of the excessive control of American education now resting in the hands of those who manage that system. Local school boards in community after community do not fairly represent, in terms of educational background and occupational status, the real majority of our citizens. A more positive form of local citizen involvement must be found.

The third barrier lies in the overutilization of *time*, and the underutilization of *performance*, as a criterion for accomplishment in American education. Educational accomplishment must come to represent more than perseverance. Some form of performance goals needs to be established and applied. Whether current approaches to performance evaluation are adequate is a moot point. Once such goals are established, the "open entry-open exit" school system combining education, work experience, work-study, and full-time employment can be formulated in terms that will make the concept of the "school dropout" obsolete and that will allow the meaningful integration of vocational education in the total school curriculum.

The fourth barrier lies in continuing reluctance to involve students in educational decision making. Students cannot forever be expected to remain as disinterested parties while others determine their educational goals and assess their educational accomplishments. The joint establishment, by students and teachers, of differing educational goals having differing standards of attainment and differing times required for completion for differing individuals must become a reality in American education.

The fifth barrier lies in continuing acceptance of an assumption that American elementary and secondary education must take place, for most students, during only nine months in any given calendar year and only in a school building. American elementary and secondary education must become a twelve months' operation and must take place in the occupational world as well as in the school itself. Eliminating this nine-months' classroom barrier would provide ample opportunities for combining education as preparation for employment with employment itself, for meaningful exchanges between educators and business-labor-industrial personnel, and for utilizing the total community as part of the educational system.

The sixth barrier lies in continuing general acceptance of the false notion that public education exists primarily for the benefit of youth. Adult education—including general education, recreational education, and occupational education—must become an integral part and a major responsibility of American public education. This approach demands that our schools be open at least eighteen hours a day six days a week and fifty weeks per year. There is no reason to believe that persons of differing ages cannot learn together, and good reason to believe that they can. If the schools can ready youth for today's occupational world, there is no reason why they should not be simultaneously involved in retraining adults for that same world.

Career education: a solution

In January 1971, USOE Commissioner Sidney P. Marland, Jr., declared his commitment to career education as a major goal of the United States Office of Education (Marland, 1971). While presenting a concept emphasizing education as preparation for making a living, Dr. Marland did not then, and has not since, formulated a precise definition of "career education." Instead, he has wisely chosen to let the definition grow from the accumulation of concerns and thoughts expressed since that time. Currently, it seems safe to say that "career education" can perhaps most accurately be described as a concept in search of a definition.

In spite of this lack of precise definition, it seems evident that career education is being viewed by the United States Office of Education as a prime, and perhaps the single most important, solution to correcting the current ills of public education. For purposes of this discussion, *"career education" is defined as the total effort of the school and the community to help all students become familiar with the values of a work-oriented society, to integrate such values in their personal value systems, and to implement such values in their lives.* Inherent in this definition is an assumption that the term "work values" encompasses a variety of work motivations including, but not limited to, those commonly accepted as contained in the Protestant work ethic. The success of career education is seen as dependent upon the quality and quantity of five components, each of which must be closely integrated with the other four. These five components include:

1. Efforts of all classroom teachers at all levels of education to emphasize the career implications of the substantive content they seek to help students learn.
2. Vocational skill training that will provide students with competencies required for successful entry (or re-entry) into the occupational world.
3. Efforts of business, labor, and industry to contribute actively to the goals of career education including the provision of work observation, work experience, and work-study opportunities for students *and for those who educate students.*
4. Comprehensive programs of career development aimed at helping students in the decision-making process while protecting individual freedom of choice.
5. Efforts to recognize and capitalize on the significant ways in which the home and family structure serve as an influence on and are influenced by the occupational society.

It is obvious that vocational skill training (or what has traditionally been called "vocational and technical education") is only one of five components in this definition. It is equally obvious that it is of pivotal importance if the goals of career education are to be attained. With this background, we now turn to specific consideration of the topic assigned to this presentation.

Broadening the base of vocational and technical training

The goal of vocational and technical training is to maximize the quality, appropriateness, variety, and levels of vocational skill training from which the individual can choose, and to provide him with competencies associated with his choices. Formal vocational education seeks to provide the individual with entry level job skills under an assumption that further expertise will be acquired during the course of employment.

Rudimentary beginnings aimed at occupational awareness and its pertinence for the individual are seen as taking place during the elementary school years. The junior high school years are aimed primarily at occupational exploration through exposure of students to basic skills required in a variety of broad occupational areas. The senior high school years consist of continuing exploratory vocational skill training for some students and more specific skill training for others. At the post-high school level, prime attention is devoted to specific vocational skill training.

Equal importance is accorded the specific goals of: (a) providing students with exploratory experiences that will broaden their basis for occupational decision making; (b) providing students with basic vocational skills having applicability in a number of different occupations; and (c) providing students with specific vocational skills required for particular occupations. The quality of a given program of vocational education is properly judged in terms of the appropriateness of balance given these three specific goals at various points in the total educational experience of students.

Measurement questions abound for students and prospective students of vocational education, for those who seek to educate such students, and for those considering employing them following their training. Examples of such questions include:

1. How can a prospective student know what type of vocational education he wants? How can he know where he is likely to be most successful?

2. How can the educational institution most effectively select prospective vocational education students? How can the school know when a student has acquired a specific vocational competency?
3. How can the competency of a vocational educator be assessed by means other than the counting of college credits?
4. How can the employer know the kinds of skills he is securing when he employs graduates of vocational education programs?
5. How can a student secure credit for vocational training gained at one level when he seeks higher level training at another level?

It seems obvious that much remains to be accomplished in the measurement field if valid answers to these and related questions are to be found. Only rudimentary beginnings have been made to date. At this point, it seems appropriate to mention a few examples of where we are as indicative of how far we have to go. No attempt will be made to provide a comprehensive review of the literature. Instead, primary attention will be directed toward areas in which work is needed.

One obvious needed area of measurement is to be found in the construction, validation, and use of simulated work experiences as a basis for vocational education decision making. The Program of Education and Career Exploration (PECE) in Georgia is one example of work currently underway in this area (Bottoms & Matheny, 1969). In this program, students are exposed to either real or simulated work experience in various occupational families which they later discuss as a basis for educational-vocational decision making. As presently constituted, it is much more a guidance than an assessment program. Self-administered career kits based on a work simulation approach have been developed by Krumboltz and his associates at Stanford (Krumboltz & Bergland, 1969), but these, too, are reported for use as guidance rather than as assessment devices. Work simulation, especially for those referred to as culturally disadvantaged, would seem to be a very appropriate assessment approach for those in the measurement field. It is not highly developed at the present time.

Fairly clear evidence exists indicating lack of utility of traditional measures of vocational interest for differentiating among students in various vocational and technical education programs (Taylor & Hecker, 1967). Still, some promising new vocational interest measures, such as the *Ohio Vocational Interest Survey* and the *Minnesota Voca-*

tional Interest Inventory are currently gaining in popularity and use in vocational counseling. The need for valid measures of work values, as opposed to vocational interests, is very obvious. That need has not been well met to date. Both interest and work value measures are needed in vocational counseling. One would hope that some use could also be made of these kinds of instruments in student selection, but this appears doubtful at present.

Considerable progress has been made in development of occupational competency testing procedures for teachers and prospective teachers of vocational and technical education (Panitz & Olivo, 1971). It appears that the National Occupational Competency Testing Project is well on its way to providing this kind of measure for those who aspire to teach in vocational education. Yet, much remains to be done to develop and validate occupational competency tests for students who graduate from formal programs of vocational and technical education. It seems strange that more work is not being undertaken in this area.

The American College Testing Program has recently initiated a new service, the Career Planning Program, specifically designed to aid students considering vocational-technical education at the post-high school level. Preliminary validity data reported in the 1970-71 *Handbook for ACT Career Planning Profile* are discouraging (American College Testing Program, 1970). The Comparative Guidance and Placement Program of the College Entrance Examination Board has reported moderately high validity figures for students enrolled in vocational-technical education program areas, but not for specific programs (College Entrance Examination Board, 1970). Certainly, it would seem that much work remains to be done in predicting student academic success in vocational and technical education programs. Again, it is encouraging to see guidance measures being developed, but discouraging to see our relative lack of selection and placement measures. The discrepancy in sophistication of measurement devices for prospective college students as opposed to prospective vocational education students is obviously great.

A concluding statement

The need for change in American public education is obvious. It now seems clear that career education will become a major vehicle for action in the total arena of needed change. As the goals and objectives of career education are emphasized, great perspective will be required in order to assure that all other bona fide goals of public education are accorded proper attention. No long-run progress will result if career education is championed at the expense of other worthwhile educational goals. Career education, if properly focused and emphasized, holds great hope for attaining the integration of vocational education into the total education program.

As progress occurs in career education, the need for assessment devices for use in guidance, selection, and placement of students will increase markedly. Appropriate measurement devices are in very short supply at present. The expertise of the measurement community is badly needed for career education to fulfill its mission.

REFERENCES

American College Testing Program. *Handbook for ACT Career Planning Profile, 1970-71 edition.* Iowa City: The American College Testing Program, 1970.

American Vocational Association. *AVA Member-Gram.* Vol. 2, No. 3. Washington, D. C.: American Vocational Association, 1971.

Bottoms, G., & Matheny, K. B. *A guide for the development, implementation, and administration of exemplary programs and projects in vocational education.* Final Report, Project OEG-0-9-207008-2779. Atlanta: Georgia State Department of Education, 1969.

College Entrance Examination Board. Comparative Guidance and Placement Program. *Interpretive Manual for Counselors, Administrators, and Faculty: August 1970 Supplement.* New York: College Entrance Examination Board, 1970.

Collins, C. C. *Junior college student personnel programs: what they are and what they should be.* Washington, D. C.: American Association of Junior Colleges, 1967.

Krumboltz, J. E., & Bergland, B. Experiencing work almost like it is. *Educational Technology*, 1969, 9, 47-49.

Marland, S. P., Jr. Career education now. Address presented at the National Association of Secondary School Principals Convention, Houston, Texas, January 23, 1971.

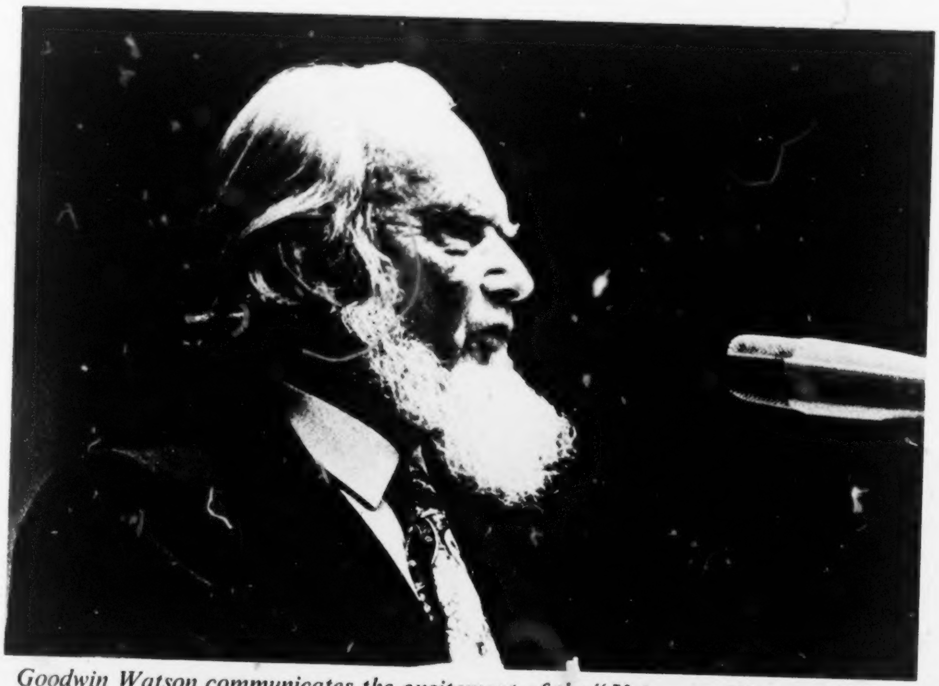
National Advisory Council on Vocational Education. *Fifth report.* Washington, D. C.: National Advisory Council on Vocational Education, June 21, 1971.

Panitz, A., & Olivo, C. T. *Handbook for developing and administering occupational competency testing.* OE Research Project No. 8-0474. New Brunswick, N. J.: Rutgers, The State University, 1971.

Taylor, R., & Hecker, D. L. *Interest and intellectual indices related to successful and non-successful male college students in technical and associate degree programs.* Final Report, OE Project No. 051-65. Big Rapids, Michigan: Ferris State College, 1967.

U. S. Bureau of the Census. *Educational attainment: March, 1970.* Current Population Reports, Series P-20, No. 207. Washington, D. C.: U. S. Government Printing Office, 1970.

U. S. Department of Labor. *U. S. manpower in the 1970's: opportunity and challenge.* Washington, D. C.: U. S. Government Printing Office, 1971.



Goodwin Watson communicates the excitement of the "University Without Walls."

Dwight W. Allen reviews Session III and offers new challenges to measurement.

Enriching Education Through Schools Without Walls

Goodwin Watson
Union for Experimenting Colleges and Universities

It is the way of revolutions to come upon us unexpectedly, even though historians may look back and see their long deep roots. The new order, in education or elsewhere, does not arise full-blown like Athena emerging from the head of Zeus. Schools without walls may be seen now as radical and revolutionary, but they have been on the way for a long time. Dissatisfactions with formal school, college, and graduate school requirements are endemic. Reforms and experiments have been trying new alternatives. Good teachers have supplemented their own resources with slides, films, tapes, and talks by visitors. Classes have explored nature and society beyond classroom walls.

Students alone, or in groups, have undertaken independent study and field projects in this country and abroad. Social action movements have enlisted student support for activities both on and off the campus.

The theory underlying a school without walls is so familiar as to be almost old-fashioned. We have known about individual differences at least since the days of Francis Galton and E. L. Thorndike. We have learned from Dewey and Kilpatrick and many others of the advantages associated with purposeful endeavor by pupils. Our own experiences have pointed to ways of learning which surpass the prescription of lesson-plans.

We have long known that children in the early years of life learn rapidly with no prescribed curriculum, no classrooms, and no certified teachers. Most of us could tell memorable learning experiences in which we have engaged on our own responsibility since leaving school. When we set our own goals, choose appropriate means of progress, and get the thrill of newly achieved knowledge, insights, and skills, we maximize psychic rewards and increase the likelihood of further learning projects.

Last month, in contrast, I listened to some 25 adults, averaging about 30 years of age, reminiscing about their school experiences.

They spoke of humiliating incidents, grisly teachers, potent injustice, boredom, repression, and miniature rebellions. Although pressed to do so by the discussion leader, none could recall a peak experience of delightful learning, during their many hours in classrooms. If this is at all typical of what laymen outside the teaching profession remember from their school days (and Jersild collected similar observations for us more than a generation ago) it is almost incredible that schools without walls have been so late in emerging.

My recent experience has been limited to college and university programs without walls, but I believe that much of what is now being tried in higher education parallels current experiments with younger age groups. I shall speak here of the developments I know best, but remind you that successful schools without walls are emerging also for younger adolescents and for children. The Parkway School in Philadelphia is probably the best known, but previous presentations on this program have referred to many other exciting developments.

The *University Without Walls* is sponsored by the Union for Experimenting Colleges and Universities of which Dr. Samuel Baskin is the president. The design for the UWW emerged from a number of college instructors meeting in a special interest group, during *Project Changeover*, an earlier project of the Union. The group set out to explore our "Dream College." What would be the most Utopian institution of higher education we could imagine? Could we integrate in one unified project the many promising trends and experiments which had emerged on many campuses? The resulting proposal was reviewed critically by distinguished educators, experienced faculty and administrators, and creative students before it was submitted to the U. S. Office of Education and the Ford Foundation which have provided funds for a Developmental Year of the *University Without Walls.* Plans have been worked out by more than 20 colleges and universities, and most of these already have students at work in an alternative undergraduate program. These UWW units have stimulated, challenged, and enriched higher education along at least eight dimensions.

New types of students

Some units of UWW are admitting students, experimentally, after only three years of high school. Most have moved beyond the usual undergraduate age range to include men and women in middle-life who desire further education but who have not been free to undertake full-time residential programs. Some recently enrolled students are in their seventies. The programs have reached out also to students who for various other reasons have not been able to use established college programs. Some are in hospitals, some in prison. Many students are financially unable to attend college unless they can also be earning a living. Some come from the minority groups which have experienced the handicaps of discrimination in our society. Many are bright students who have tried the standard undergraduate offering and have become disillusioned and alienated. A few demonstrate the kind of genius which warrants and demands opportunity to grow in rather narrow and specialized channels. Others earnestly seek an education more useful than present college offerings in facilitating their action for social change.

Such diversity inevitably enriches peer-interactions beyond what is now available in most colleges. While periods of campus residence may be brief, the interaction among students will continue in settings near their homes and work, as well as at field centers. The new varieties of students, with their many distinctive needs and attitudes, have led to other changes that also enrich education.

New content of curricula

The program of each student in the UWW has been devised especially and uniquely for him. He will have proposed projects which interest him and promise to enrich his life. He has responded to ideas from his advisers and from fellow students. His targets may include some which have commonly been stated objectives in undergraduate education (he may want to learn Spanish or anthropology), but they are very likely to include many beyond the range of present college offerings. One important expectation is that each student, as he records and evaluates what he had done, will grow in his competence to select and to invent more rewarding educative activities to meet his own needs and values. Many of these innovative enterprises will transcend the contents of even the richest college catalogs.

New learning resources

Accompanying the extended curriculum will be an enlarged range of learning resources. When the walls of college classrooms, libraries, laboratories, studios, and dormitories no longer limit students, excellent resources to aid learning are likely to be found also out in the world beyond the campus. A memorable statement by John Dewey said: "The most Utopian thing about Utopias is that they need no schools." Insofar as the society has developed satisfactory ways of feeding, clothing, housing, employing, and governing its citizens, along with provisions for health and the enjoyment of life, these social processes can convey to students at first-hand an understanding of their culture, economics, politics, technology, and arts. Insofar as the society is defective—as for example, in its urban ghettos or polluted environment—no pamphlets or films can drive home the need for change as effectively as do actual experiences while living in the grim reality. The best theatres, symphony orchestras, farms, mines, factories, banks, art collections, architectural achievements, airplanes, baseball and football teams to be seen today are not on college campuses. Students are right to prefer direct encounters to second-hand reports.

New faculty roles

In the UWW a teacher is a diagnostician of learning readiness, a guide to an immense range of potential resources, a listener and facilitator, a critic and evaluator. He helps students set their own goals, choose means for attaining them, and gives feedback on progress toward the student's own goals. This is a different task from that of measuring achievement by assessing progress toward the teacher's prescribed goal. These new tasks are likely to call for talents the faculty might never have discovered had they continued with only their earlier teaching materials, methods, and habits. Teachers will still conduct orientation groups on learning skills and often meet with seminars in the field, but much of their time and ingenuity will go into new roles.

New supplementary staff

If students are to do much of their learning off-campus, in connection with their employment or leisure interests, or service to others, their effective instructors will be their supervisors on the job, their companions in recreation, performers in movies or television, and community leaders. If they venture into subcultures very different from those they have lived in at home or in college, almost everyone they encounter may become a source of new learning. If they seek out experts on particular problems, areas of knowledge or skills, these become their advisers. We are using the term *Adjunct Professor* to designate these abundant new resources in higher education. For many students, practitioners have higher credibility than do professors. "Those who can, do! Those who can't, teach about doing." Ivan Illich in a recent article (1971) comments: "The best teacher of any skill is usually someone who is engaged in its useful exercise [p. 12]."

A word might be said also about the enrichment the Adjunct Professors will find for themselves in their new pedagogical roles. In groups of businessmen, I have sometimes inquired, "If you were guaranteed your present income, regardless of what work you did, what would you choose to do?" The most frequent answer was that the men would continue in their present jobs, but second in frequency was the response: "I think I would like to teach." For many competent workers, a little bored with their job routines, the chance to interact with an interested and curious apprentice will provide a new dimension of experience. Both master and apprentice will learn from each other.

New ways of communicating

The endeavor to continue a dialogue between an adviser who lives in a college community and a score of advisees scattered across the nation or the globe, taxes ingenuity. Telephone calls and conference calls may become more frequent than class meetings. Audiotapes and kinescopes may supplement written reports and term papers. The computer may become preferred to library and catalogues as a way of retrieving desired facts. Television has already shown, in some BBC Open University programs for example, educative possibilities surpassing most textbooks.

New assessment procedures

Most college students get very little feedback on their personal growth except in the one dimension of answering test questions or writing papers about assigned course content. This is a rather narrow and barren dimension of achievement even for the small minority who go on to graduate degrees and college teaching or research. The UWW, with its new curriculum units, its immense range of learning resources and faculty, with its emphasis on attention to the whole person of each individual, enriches the conception of education for personal growth. Students become aware, not primarily of scores on achievement tests or grades in courses, but of their status and growth in intellectual curiosity, fairmindedness, imagination, resourcefulness, social concern, aesthetic judgment, and ability to design a life which, as prescribed by Socrates, has been critically examined. As the measures become more diverse and representative of human potentialities, students pay attention to more kinds of growth-needs and education is correspondingly enriched. It is almost a truism today to note that what we measure and give credits for tends to become what we teach.

The various units of the UWW have already agreed on some basic principles, among them:

a. Competence of the candidate rather than courses or credits shall be the criterion.

b. Readiness for a degree includes concern for both processes and products; both cognitive and affective growth; both breadth and depth; both what a student has acquired and what he has contributed.

c. Each student keeps a cumulative record—a log marked off in learning episodes—each evaluated as a learning experience.

d. The log is supplemented by external or objective appraisals from fellow students, faculty advisers, adjunct professors, field observers, and other informed sources.

Goodwin Watson

New ways of administration

The units of the University Without Walls are demonstrating a new model in college administration. There are still admission functions but they may soon expire. There may be no registrar other than a machine which can record and duplicate, on demand, the transcript of learning experiences prepared by a student and his advisor. No dormitories are needed except for brief periods of residence. No student personnel functions remain, over and above the work of faculty advisers. There will still remain a finance office, which receives and disburses funds. Most of the maintenance costs and staffs dwindle as students plan to do most of their work off-campus. The net result will almost certainly be a reduction of cost per student below recent astronomic figures.

A major change is that from a college run by the administration and/or the faculty to a program in which students share in all vital functions. Students and staff together decide on core faculty and their salaries, adjunct professors and their compensation, tuition and other budget items, educational assessment procedures, graduation standards and their application in individual cases. New patterns of decision making emerge when students participate genuinely and not merely in a token fashion.

We have seen that schools without walls have enriched educational offerings by bringing in new types of students, new curriculum content, new aids to learning, altered faculty roles, new sources for teachers, new ways of communicating, new assessment procedures and administrative operations which are less costly but which utilize student contributions more fully.

What, now, are implications for the measurement movement?

The implications are truly revolutionary! Most agencies publishing standard tests have assumed some agreed-on content to be measured. With multifarious students, having a wide variety of objectives, utilizing numerous educational procedures to achieve individually unique goals, the whole structure of the commercial test-publishing business seems to be threatened.

We do foresee a need for many self-administered tests which help a student decide on appropriate educational goals for himself! These will develop rather than disrupt the relationship between the student and his faculty adviser. They will leave the initiative for test use with the student rather than the faculty. They will be primarily diagnostic rather than judgmental. They will relate to the affective and conative dimensions of growth as well as to the cognitive. There may be a few tests today which range over all three aspects of growth and are designed to encourage better learning rather than to provide credits to students by outworn and irrelevant assumptions. We are likely to need many more.

The measurement movement will have come to a drastic turning point if the assumptions of the University Without Walls are gradually adopted by most elementary, secondary, undergraduate, and graduate schools. Competence will then be seen to be as multidimensional as music, art, or personality. Tests will then be sold to learners rather than to teachers, to facilitate rather than to pass judgment on self-directed learning. Most of the statistical indices customarily applied to evaluate tests will have become irrelevant. The integration in the newly emerging schools without walls, of many familiar truths about learning is bringing changes so far-reaching as to be, indeed, an educational revolution.

REFERENCE

Illich, I. After deschooling, what? *Social Policy*, 1971, **1**, (3), 5-13.

Discussion

DWIGHT W. ALLEN
University of Massachusetts

To evaluate change phenomena in not one but three active educational areas is no mean task, but let's see what we can do.

First of all, when you look at early childhood education, you confront evidence of many *competing* needs. There is evidence that I.Q. differences are *not* registered at early ages; there is evidence that, even in the worst ghetto environment, a child has a better chance for survival than in a sterile orphan's environment, and in fact, there are a host of competing measurement needs that lead one even to look at early childhood education.

Second, if you look at career education, you see a complete lack of access for the very people who need access to career education; you see the way in which society stamps out the identity of the individual, depending on his career, but you also see "career" in general becoming less and less a major part of people's life space. Yet the "career identity" survives. We still have doctors, lawyers, teachers, or bricklayers, even though careers are becoming a smaller part of our lives. The question is how to devise educational strategy that will make some sense out of these conflicting factors.

Third, there are the nontraditional educational approaches that are being proposed. Open admission, for example, is proposed as an alternative to racially and socioeconomically prejudiced admission standards; but in a very subtle way that proposal may suggest that the only option to racist standards is no standards at all. And, it seems to me that in many ways the "no standards" position is a very racist position. In the quest for standards we may find ourselves abrogating any standards, and thus perpetuating our racist tendencies.

I would like to propose that we establish an educational system where "lower" education is from ages 3 to 13 years. This would restore the ending point for lower education to approximately the age of puberty. At the time that the original separation of "lower" education and "secondary" education was established, 18 was considered the age of puberty. So now with "puberty" coming at an earlier time in a child's life, we have three or four years of what amounts to "prepubescent education" coming *after* puberty.

So I would propose an educational system where "lower" education is from ages 3 to 13; where "higher" education is from ages 13 to 16; where at 17 everyone would go off for two years of national or state service to humanity; and where ages 18 to 21 can be for post-graduate education. I think that that would finally place nontraditional and vocational education in one package.

Let's look at the measurement implications of such a proposal. Initially, there is the very simple question of what evidence is necessary for decision making about such a program. I maintain that if we approach this question with the traditional measurement strategies, it will be about the year 3000 before we have enough information to make any decisions whatsoever.

The challenge is: How do we refocus measurement strategies so that they become realistic aids to decision making? How do we refocus measurement strategies so that we can demand of those strategies some practical help in the real-time decision-making process?

One thing needed is to require measurement people to take more responsibility for helping decision makers, rather than having the measurement experts dream off in a vacuum somewhere, while the decision makers are off making decisions in a vacuum, without *any* data. Maybe decision making should be tied to a "reasonable level" of evidence. The decision makers should ask: What measurement evidence is available? And what data that I need can be obtained in time for me to make a decision? And then, equipped with the answers, they should make the best decisions they can—given at least *some* relevant information.

One thing that concerns me is that I see little evidence that educational research in its current form is going to make a difference to educational practice. We have lots of research evidence, for example, that girls should start school a year earlier than boys, yet I fail to see this finding reflected in any educational programs. Instead, what I see happening is measurement being used as an excuse to block educational innovation. This situation may not be the intent of measurement people, but it *is* the work of many decision makers who use data inappropriately. Thus, in the name of the measurement specialist, innovations are often destroyed.

I agree with what Lee Cronbach said this noon—that *who* asks the questions is crucial. But I'm not sure that I would agree that the measurement specialists ask the questions most important for educational decision making. I agree, however, that they have indulged in the luxury of asking the questions they can most easily answer. Unfortunately, these questions do not always turn out to be of much practical value. They turn out to be questions that are easily investigated and neatly oriented to prevailing psychological models, but they often are of no value to the making of practical educational decisions.

I would rather turn the process upside down. I would rather be the one, as a practitioner faced with real decisions, to tell measurement specialists what problems to investigate, and what questions I need answered to make an intelligent decision. Give me all the provisos—all the "ifs," "ands," and "buts" about how I ought to use the tentative and early findings you have. But also, most importantly, please give me the benefit of your best judgment on the pressing problems that I *must* deal with.

In a society characterized by change, education has remained immobilized and is rapidly becoming an anachronistic vestige of the nineteenth century. Education is more likely to fail as a result of our uncertainty—because of inaction that is justified by a call for more evidence—than it is likely to fail because of radical experimentation.

Is measurement a contribution when it is happily remote from the practical problems of educational practice? I think not. Measurement must suggest intervention strategies that are not only responsible in reacting, but that may directly influence the process of educational change. The role of objectivity must be seen in a new light. The measurement person must be seen as a responsible advocate. Unfortunately, those researchers who define the parameters of education are not involved in its practical use. Let me illustrate this with a problem that recently involved me.

Harvey Scribner and I were talking about the problem of getting high school diplomas for more of the kids in New York City—the kids who have been systematically dumped, largely on racial grounds. The problem is that no one knows what "education" is. No one really knows what it means to be "educated" in a society that has gone whirling off in several different directions simultaneously—where, for example, access to informal education surpasses access to formal education.

So, Harvey and I dreamed up a system. Why not give high school diplomas to all students who can hold down a job that usually requires a high school diploma? Or to students who have satisfactorily completed a year of college? A student who completes a year of college obviously has satisfied high school requirements. And then, to put a little incentive into the system, let us say you give the employer or the college a $1,000 bonus for every student it services with a high school diploma. Now, the problem is that, rather than giving me data on questions *I* need answered to implement this proposal, researchers will typically cite all the old data showing how my proposal violates psychological theories. But what I really need are practical answers to practical questions.

I think that we need to address ourselves in education to problems of equity, not equality. We are concerned far too much with equality as a part of our sacred trust in "objectivity." This tradition of the objectivity of education can be really insidious. It is characterized, to start with, by the fact that the ideal teacher must be neuter. Or, if he cannot be neuter, he must be neutral; and if he cannot be neutral, he must at least pretend to be neutral. I can think of few things worse than a biased teacher pretending to be neutral and sneaking up on kids, but that is the tradition. It is the tradition of objectivity, and a tradition that gets reflected in our notions of equality.

You put Mohammed Ali in the boxing ring with someone who is starving to death and say: Okay, as a matter of equality you are going to fight by the same rules. And if the floor gets wiped up by the guy, you pick him off the floor and say: Want to go another round? The rules are the same for everyone—everybody is equal.

The notion of equality in a society characterized by inequality can be an unequal, insidious, inequitable, and racist notion. Somehow we must develop strategies that cater to *equity* rather than to a false notion of equality—strategies that recognize existing inequalities and recognize that an important objective is to overcome these inequalities—which means that sometimes you do not apologize for unequal treatment. You may, in fact, *seek* unequal treatment in the name of equity.

We have to change our concepts of measurement. We have always recognized differences based on measurement. But we are just now beginning to see the differential consequences of these differences. Somehow, the measurement profession has to take responsibility for the consequences of the differences that it verifies—the recognition of difference, the access to differences, and the consequences of a philosophy of differences.

Frequently, when we cannot measure as much as we would like to—and thus find ourselves measuring some baseless abstraction—the abstraction is not based upon the things that are most important in practice, but upon the things that are most easily measured. Then, we are forced to build educational programs based on a lack of evidence.

For example, in a collegiate program the only thing that is measured is course performance. However, there are many people who would argue that the most important part of a college education is the living environment, the association with peers outside the classroom. But no one has yet suggested ways in which effects of the out-of-class environment can be measured. Thus, we perpetuate a system over the centuries that measures and is based on only a small segment of the reality of students' experience. And, at the same time, we want to determine the equivalency of some new programs, like the University Without Walls, by trying to relate their equivalency to the narrow and very coarse band of evidence that measurement and convenience have selected out of the totality of student experience.

One big question I would ask about the University Without Walls is this: If you could put together a program that would include the entire course part of a college degree, but without the enriched personal associations, where would you be? I happen to believe, as an act of faith, that the experiential kind of learning that might be substituted for traditional coursework—the breadth of experience of, say, having supper in Istanbul reflected for academic credit—is probably as valuable, if not more so, than the coursework. But how then are we to get usable evidence on such questions unless researchers are willing to supply at least tentative data on the practical, and messy, problems of educational practice?

I can perceive at least eight problems in the development of measurement strategy for new directions in education. As I said before, the first is that no one really knows what "education" is, independent of whatever happens under current arrangements and by current measurement strategies.

Second, the relationship between knowledge and experience is unclear. Thus, although the introduction of experiential elements is a positive achievement, it is not clear what percentage of education is appropriate for that element.

Third, the assumption that education is linear is untrue; and yet all of our evaluation procedures are oriented to the assumption of linearity.

Fourth, because of the vast increase in knowledge and specialization, because of random access to this knowledge, and because of the increased diversity of current curriculums, degrees themselves do not mean what they used to mean. It is becoming less valid—if not completely invalid—to compare two degrees in the same program in different institutions, or even to compare degrees in the same program in the same institution at different times. Where do you go then in terms of trying to figure out what "degrees" are?

Fifth, monitoring the new, diverse activities to which students at all levels are being directed presents, at best, organizational uncertainties.

Sixth, there is great pressure to compare new programs with traditional programs. Such comparison is often totally inappropriate because the social context of education is changing, and its roles and functions will continue to change qualitatively and quantitatively. Educational programs must change correspondingly, in objectives as well as in instruction and content.

Seventh, the administrative problems of coordinating the new resources available to students are considerable. Once the student enters a certain network of resources, then he must maneuver through the network. Entering the network is the first problem, after which the student can assist other students in gaining access to the network. Some students become a part of the problem and a part of the solution at the same time.

How are we going to deal with these problems in the measurement context? The problems of designing measurement strategies to determine the adequacy of communication of the resources available—let alone to evaluate the resources themselves—are formidable.

Finally, one of the central problems in developing innovative systems is that there seems to be a critical point in time where the program "clicks" and the spirit of innovative thrust is embodied in the program. Before this point is reached, the program cannot be called a success. In order to understand this point, wise judgment must be used, and often dynamic and inspirational leadership is needed. But it is important to keep in mind that the process of reaching the critical mass is often whimsical, and thus not conducive to empirical evaluation. Intuition, judgment, and vision are at a premium in the development of these programs.

Dwight W. Allen

We need to develop basic changes in our conceptions of schools and students to make our ideas more consistent with the mobility and sophistication of the American student population, as well as with the increasing integration and interdependence of all the threads of our social fabric. Maladjustments in our educational systems are reflected in the increasing inability of our students to adjust to the responsibilities and demands of a rapidly changing society. Let's develop a sense of "the evidence" which is reasonable, and with that consciousness move forward decisively to attain that minimum level of data we need for real decision making, recognizing how tentative and uncertain even our most informed judgments might be.